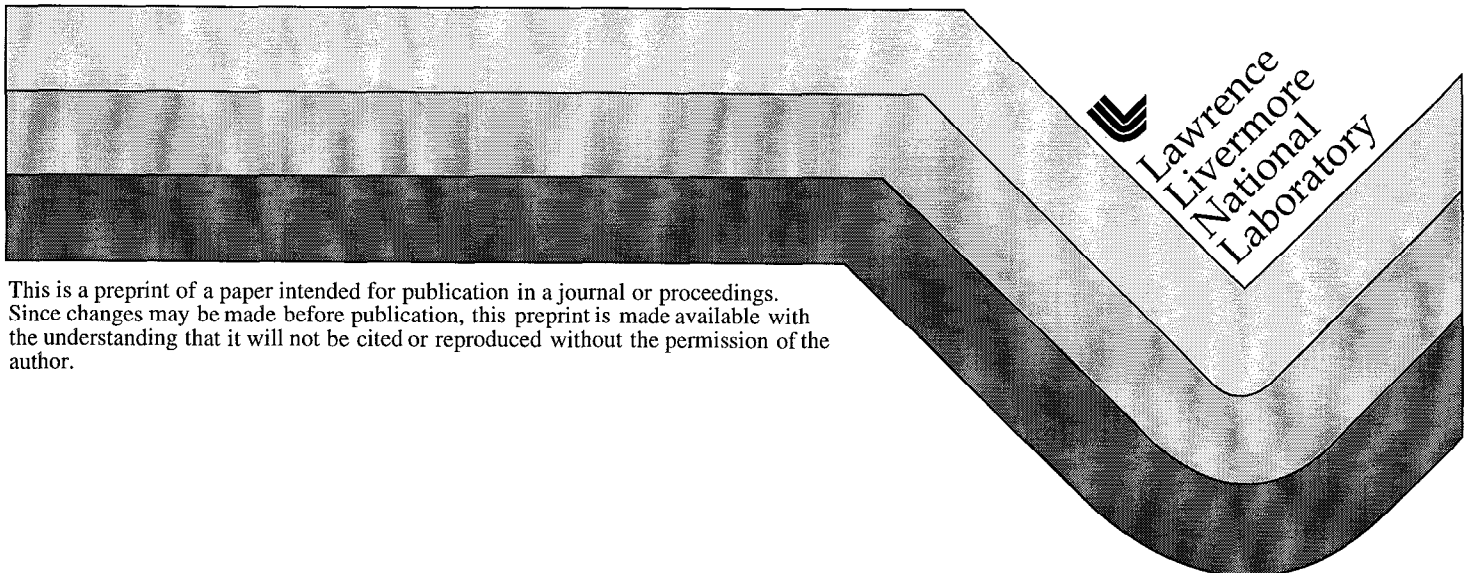# Objective Analysis of the ARM IOP Data: Method and Sensitivity

M.H. Zhang
J.L. Lin
R.T. Cederwall
J.J. Yio
S.C. Xie

This paper was prepared for submittal to the

First International Symposium on Turbulence and Shear Flow Phenomena
Santa Barbara, CA
September 12-15, 1999

April 1, 1999

DISCLAIMER

Objective Analysis of the ARM IOP Data:

Method and Sensitivity

M. H. Zhang, J. L. Lin

State University of New York at Stony Brook

R. T. Cederwall, J. J. Yio, S. C. Xie

Lawrence Livermore National Laboratory

ABSTRACT

Motivated by the need of to obtain accurate objective analysis of field experimental data to force physical parameterizations in numerical models, this paper first reviews the existing objective analysis methods and interpolation schemes that are used to derive atmospheric wind divergence, vertical velocity, and advective tendencies. Advantages and disadvantages of each method are discussed. It is shown that considerable uncertainties in the analyzed products can result from the use of different analysis schemes and even more from different implementations of a particular scheme. The paper then describes a hybrid approach to combine the strengths of the regular grid method and the line-integral method, together with a variational constraining procedure for the analysis of field experimental data. In addition to the use of upper air data, measurements at the surface and at the top-of-the-atmosphere are used to constrain the upper air analysis to conserve column-integrated mass, water, energy, and momentum. Analyses are shown for measurements taken in the Atmospheric Radiation Measurement Programs (ARM) July 1995 Intensive Observational Period (IOP). Sensitivity experiments are carried out to test the robustness of the analyzed data and to reveal the uncertainties in the analysis. It is shown that the variational constraining process significantly reduces the sensitivity of the final data products.

1. Introduction

Deriving atmospheric wind divergence, vertical velocity, and advective tendencies from field experimental data is an old topic. It gained renewed attention in recent years because of the need to use these derived quantities to force and evaluate physical parameterizations in General Circulation Models (GCMs) (Randall et al. 1996). Two recent major field experiments were specifically set up for the purpose of forcing model parameterizations with observed data. These are the Tropical-Ocean and Global Atmosphere's-Coupled Ocean and Atmospheric Response Experiment (TOGA-COARE) and the ARM Program of the Department of Energy (Webster and Lukas 1992; Stokes and Schwartz 1994). Most parameterized processes, such as cumulus convection and clouds, occur on time scales much less than a day. In order to meaningfully interpret model results of these processes against observations, accurate forcing data of high temporal resolution are needed. This poses significant challenges to the objective analysis of field data.

The main problem in deriving accurate objective analysis from a field experiment is the insufficient sampling of measurements, attributed not only to invalid or missing measurements, but also to scale aliasings. The actual measurements contain information of all scales including those that the observational network cannot resolve. Thus, it is not always appropriate to directly use point measurements to extract area-mean quantities. Limitation in the accuracy of instruments and measurements is another source of concern. When a field experiment is conducted in a small region, instrument and measurement errors can have a large impact on the derivative fields.

After the completion of the GATE experiment in 1969, Ooyama made a five-year concentrated effort to analyze the GATE data (Ooyama 1987). For TOGA-COARE, Lin and Johnson (1996a, 1996b) used several methods to analyze sounding measurements. They tested linear and quadratic fitting methods and concluded that the methods were not suitable due to the uneven distribution of COARE stations. They used the Barnes interpolation scheme in the final analysis (Barnes 1964). The cubic-spline mechanical fitting package from Steve Lord described in Ooyama (1987) was also tested. It was concluded that the results are similar to those from the Barnes analysis. Frank et al. (1997) also analyzed the TOGA-COARE data but used the line-integral method. Moisture budget from their analysis is very different from that in Lin and Johnson. For example, the IOP mean of diagnosed precipitation is 10.5-11.8 mm day^{-1} in Frank et al. versus 5.7-6.1 mm day^{-1} in Lin and Johnson. This magnitude of uncertainty is likely to be typical of most existing objective analyses. Uncertainties in the transient data are even larger. The impacts of possible errors in the analyzed data on the model simulations have been shown in Ghan et al. (1999) for ARM and Redelsperger

et al. (1999) for TOGA-COARE. Continued efforts are therefore required to improve these analyses.

This paper describes an objective analysis scheme for ARM. We first review several of the available methods and schemes, and highlight their advantages and disadvantages. We then introduce a hybrid approach of combining the regular-grid method and line-integral method in conjunction with a variational constraining procedure (Zhang and Lin 1997). The constraining procedure uses additional surface and TOA measurements to assure consistencies in the objective analysis. We finally carry out sensitivity experiments to examine the robustness of the final data.

The paper is organized in the following order. Section 2 reviews the commonly used objective analysis methods. Section 3 introduces the hybrid approach and describes the implementation of the variational constraining procedure for the July 1995 ARM IOP data. Section 4 presents the sensitivity study. The last section contains a summary of the paper.

2. A Review of commonly used objective analysis methods

Methodologically, horizontal derivatives, when averaged over a field experimental domain, can be calculated using one of these three methods: analytical method, line integral method, and finite difference method. The first method uses an assumed functional form of the spatial distribution of a field. The least square fitting of this function to the observed data is used to derive the free coefficients. Spatial derivatives can then be analytically evaluated from the fitting functions. We refer this type of schemes as fitting method. The second method uses line integral of fluxes passing through the boundary of a selected domain to calculate the area-averaged divergence field. Linear assumption is made in the calculation of flux through two stations (Figure 1a). The third method consists of establishing regularly spaced grids, interpolating or extrapolating observational data to these regular grids, and then taking the standard finite difference of the interpolated fields. We refer this method as regular-grid method (Figure 1b).

Technical implementation of the fitting method and the line-integral method are rather straightforward. With three stations, Davies (1995) showed that linear fitting and line-integral calculation of the divergence fields are exactly the same. With more stations, results from these two methods differ because of their very different assumptions, one being a prescribed functional form within the analysis domain, and the other being only related with observations at the domain boundary. A key element in the fitting method is the choice of the functional form, which can be

quite subjective. There is no magic prescription for this form. For example, Thompson et al. (1979) experimented both linear and quadratic fitting in their analysis of the GATE data, while Ooyama (1987) used cubic spline fitting.

For the line integral method, potential problems can arise from the relative locations of the adjoining stations. For example, if the analysis domain is triangle ABC in Figure 1a, the flux calculation through boundary AB would only use measurements from stations A and B. Wind convergence would be inferred since the wind at C is stronger than those at A and B. In reality, however, measurement at D suggests that this region is most likely experiencing a wind divergence. This type of problem is not uncommon in field experiments because of logistical reasons. Differential drifting of balloons at different stations poses a similar problem. Figure 2 shows the drifting of balloons launched at around 11:30 GMT on July 22, 1995 at the four ARM SGP balloon facilities. Typical ascending time of a balloon from the surface to 200 mb level is about 1 hour. With an averaged wind speed of 15 ms^{-1}, a balloon could have drifted over 50 km at that level. As shown in Figure 2, the change of the relative position could be as large as 0.2 degrees in latitude and longitude, which is about 20% of the distance between two balloon stations at the ground. An extreme scenario is that convergent winds move the balloons close to each other and divergent winds move the balloons away from each. Direct use of the balloon measurements in the line integral method, therefore, may introduce large errors. In section 3, we describe a method that attempts to avoid these problems.

The regular-grid method deserves more discussion, not only because it is widely used in the literature, but also because this method is also used to fill missing data in the line-integral method and in some fitting calculations. In comparison with the line-integral method, the strength of the regular-grid method is that it can easily combine measurements from various platforms and from various stations. The weakness is that potential problems in the original measurements become hidden in the analysis. The main step in the regular-grid method is to interpolate measurements to the analysis grids. The general form of the interpolation and extrapolation of a field f to a set of prescribed grids can be written as:

$$f_a(x_i) = f_b(x_i) + \sum_{k=1}^{k=K} w_{ik}[f_o(x_k) - f_b(x_k)] \tag{1}$$

In (1), subscript "a" represents analyzed field at the N analysis grids x_i with i =1,2,...N and x_i being a three or four dimensional vector (when time is used). Subscript "b" represents a background field that is used to assist the analysis. The background field can be the climatological means or outputs from operational models. If no background field is needed, f_b can be set to zero.

subscript "o" represents observational data, which are made at K stations at x_k with $k=1,2,...K$. w_{ik} is the weighting coefficient of measurement at x_k for the analysis at x_i. Sometimes the calculation from equation (1) is considered as an estimate of the background field, the final analysis can then be iteratively obtained as

$$f_a^{(n+1)}(x_i) = f_a^{(n)}(x_i) + \sum_{k=1}^{k=K} w_{ik}[f_o(x_k) - f_a^{(n)}(x_k)] \qquad (2)$$

where the superscript n represents the sequence of iteration. If the iteration is carried out up to the convergence of the solution, the second term on the right of equation (2) becomes zero, and the analyzed data matches observations at the measurement grids. This is often an undesirable feature because the true spatial variability of a field is never the same as what is implied by the spatial structures of the weighting coefficients and measurement errors should be considered. The most widely used interpolation schemes in the literature are the Barnes scheme (Barnes 1964), the Cressman scheme (Cressman 1959), and the statistical interpolation scheme (e.g., Ooyama 1987).

a. The Barnes scheme

The Barnes scheme does not use a background field. The initial analysis is simply a weighting average of existing measurements. The weighting coefficients are calculated based on the distance between an analysis grid and the measurement stations:

$$w_{ik} = w(x_i, x_k) = w_o \exp[-\frac{(x_i - x_k)^2}{L^2}] \qquad (3)$$

with $\sum_k w_{ik} = 1$. All measurements are used for the analysis at each grid. This is both a blessing and a drawback. The blessing aspect is that the calculation can be carried out regardless of the amount of available measurements. The pitfall is that one or two available measurements can fill all grids in the whole analysis domain. In the Barnes scheme, the weighting of observational data is blind to the actual location of the observational station other than the distance. As a consequence, if two stations are very close to each other, such as overlapping balloon sounding and profiler stations at the same location, measurements from that location would carry double weight.

To demonstrate the impact of the free parameters in the Barnes scheme on the analysis, namely the length scale L and the number of iterations, we use the analysis of the eastward component of the

500 wind at 11:30 GMT July 22, 1995 at the ARM SGP site as an example. The analysis uses all measurements from five balloons and 17 profilers at 500 mb as input. Figure 3a shows the Barnes analysis from the first iteration with length scale L = 50 km. The locations of the five ARM balloon facilities are labeled in the figure. The local impact of data at the B1 facility can be seen clearly in the figure. Figure 3b shows the same analysis except that the length scale is changed to 200 km. With this length scale, the analysis field is drastically smoothed. With the same length scale of L=200 km, Figure 3c shows the Barnes analysis from the fourth iteration. The iteration regenerates smaller scale features. Increasing the number of iterations is somewhat equivalent to reducing the length scale. The analysis from the Barnes procedure, as what can be expected, is sensitive to the choice of the length scale and the number of iterations (Leach et al., 1996). By taking the difference of the analyzed field at two stations, say B1 and B6, one can get a measure of the difference in the derivative field among the three analyses in Figures 3a to 3c. The difference is quite significant.

b. The Cressman scheme

The Cressman scheme was originally designed to efficiently carry out operational analysis of meteorological fields in a large domain. The weighting function in Cressman (1959) took the form of

$$w_{ik} = w(x_i, x_k) = \begin{cases} \dfrac{1}{N_i} \dfrac{L^2 - (x_i - x_k)^2}{L^2 + (x_i - x_k)^2} & , \qquad d_{ik} < L \\[2em] 0 & otherwise \end{cases} \tag{4}$$

where L is a length scale, d_{ik} is the distance between an observation station and an analysis grid. Here, N_i is the number of measurements within distance L of the analysis grid x_i. A background field is required. Interpolation is carried out for the difference field between observations and the background, rather than the observation itself. The original Cressman implementation also used the geostrophic relation and winds for the analysis of the geopotential height. This is not discussed here. The weighting of measurements is restricted to within a distance L of an analysis grid. Because of this restriction, if an analysis grid is away from any measurement station farther than L, the analysis at this grid will only take the background field.

We use the mesoscale analysis from the NOAA Rapid Update cycle model (RUC) as our background. The impact of background field disappears when the iteration in (2) converges. The

profiler and sounding stations spread over the whole ARM analysis domain, this reduces the impact of the background field. Important parameters in the Cressman scheme, similar to the Barnes scheme, are the choice of the length scale L and the number of iterations. As can be expected, their impacts on the analyzed data are very similar to those in the Barnes scheme, and are thus not separated plotted here.

c. Statistical interpolation

The statistical interpolation scheme avoids the subjective specification of the weighting coefficients. The general form of the statistical interpolation can be written the same as equation (1). The time dimension is often used to form a statistical ensemble. The weighting coefficients are calculated based on the distribution of the covariance fields. They are obtained through the minimization of error variance described as:

$$\varepsilon^2 = < \sum_{i=1}^{I} [f_a(x_i) - f(x_i)]^2 >$$

(5)

where $f(x_i)$ is the "true" field at the analysis grid x_i. The bracket represents the ensemble average (time average). Without the use of a background field, the analysis is written as

$$f_a(x_i) = \sum_{k=1}^{k=K} w_{ik} f_o(x_k)$$

(6)

and the weighting coefficients are calculated as (Daley 1991):

$$\overline{\overline{M}} \, \overline{w}_i = \overline{P}_i$$

(7)

where $\overline{\overline{M}}$ is the KxK covariance matrix among observational stations, and $\overline{P}_i$ is the Kx1 covariance vector of the true field at the analysis grid x_i with all measurement stations x_k, $\overline{w}_i$ is the Kx1 vector of measurement weightings for the analysis at x_i. This procedure is expected to discount the weighting of stations that are significantly correlated with each other.

The difficult step in this analysis procedure is the calculation of the covariance field involving the true field. Some approximation of the true field is needed to obtain the covariances, which could be the analysis from the Barnes scheme, the Cressman scheme, or a background field directly. Figure 3d shows the statistical interpolation of the eastward wind component, in which the Barnes analysis of Figure 3c is used in the covariance calculation. It is seen that the analysis has not eliminated the small scale structures associated with a low-value data point at B1. Furthermore, the analysis is in fact quite sensitive to the covariance matrices, as will be further discussed below.

7

d. The Ooyama scheme

The analysis scheme used in Ooyama (1987) for GATE belongs to statistical interpolation. It is separately listed here because it specifically emphasized the de-aliasing of small-scale features in the observations. Balloon soundings are taken at a fixed location and fixed time representing more or less a point measurement. The purpose of objective analysis, however, is to use the measurements to obtain the spatial and temporal structures over the network, rather than over a small area near the balloons. The challenge of objective analysis is to discern which part of the measurement is local and which part is resolvable.

The procedure implemented in Ooyama (1987) is a modification of the standard statistical algorithm. Ooyama (1987) considered that the true field contained undesirable small-scale features. He therefore minimized the error variance of the analysis with respect to a "smoothed true field" rather than the "true field". A penalty function is used to measure the distance of the smoothed field and the true field. The cost function is written as:

$$\varepsilon^2 = < \sum_{i=1}^{I} [f_a(x_i) - \bar{f}(x_i)]^2 > + < \sum_{i=1}^{I} [\bar{f}(x_i) - f(x_i)]^2 > \tag{8}$$

where $\bar{f}$ is the "smoothed true field." The actual calculation of the weighting coefficients is exactly the same as in the standard statistical method, except that the smoothed true field is used in the covariance calculation for $\bar{P_i}$. This is because the second term in (8) does not depend on the weighting coefficients. Ooyama used a cubic spline fitting, with a penalty function to limit the magnitude of the horizontal derivatives, to obtain the smoothed true fields. Use of a smoothed field for the calculation of the covariance is one unique aspect of the Ooyama scheme.

The second unique aspect of the Ooyama scheme is to further filter available observations used in the weighted mean calculation in (6). At any given time, the spatial structure of the observations is expanded according to the Empirical Orthogonal Functions (EOFs) of the covariance matrix $\overline{\overline{M}}$, and only those components with significant contributions to the total variance of the field are retained. Ooyama (1987) showed this filtering to be also effective in desensitizing the statistical analysis.

These two aspects, one modifying the weighting coefficients, and the other modifying the weighting data, both tend to smooth the analysis field. Figure 3e shows the statistical analysis with the "true" field smoothed by a simple 9 point box averaging, which is otherwise the same as

in Figure 3g. The covariance vector $\overline{P_i}$ is therefore slightly changed. It is seen that the analysis becomes smoother, owing to the change in the weighting coefficients. Figure 3f shows an analysis that further filters the observation by using only those EOF components that explain 95% of the total variance. It is seen that this step is effective in reducing the impact of the small local feature at B1, which is filtered out in the reconstruction of EOFs.

These results demonstrate that the final data product is sensitive to the analysis scheme, and perhaps even more to the actual implementation of a particular scheme. The length scales and the iteration times in the Cressman and Barnes schemes all have large impact on the final analysis. So does the filtering of data and the calculation of the covariance matrix in the statistical analyses. Major advantages of using the Cressman and Barnes schemes are their mathematical simplicity, especially with the Barnes scheme where no background field is needed. Comparing the Barnes scheme with the Cressman scheme, if a background field can capture the general spatial and temporal structures of an analysis field, then the Cressman scheme is likely to outperform the Barnes scheme. Major weakness of both these schemes is the dependence of the calculation on the assumption of the length scale the number of iterations, and their blindness to the actual locations of the observations other than the distance. Serious problems can occur when the stations are ill positioned or when there are few measurements, and these problems are hidden to the users.

In comparison with the Barnes and the Cressman schemes, the statistical scheme is more objective in the determination of the weighting coefficients. Weakness of this scheme is the subjectivity in the estimation of the covariances, and the more complicated calculation procedure. Furthermore, a field experiment is often limited in length and the sample size used in the correlation calculation is often small. Even if the sample is long, the actual data at the analysis time may not necessarily follow what is described by the correlation patterns.

An incomplete list of previous studies that have one way or the other addressed the objective analysis of field experimental data is given in Table 1. Regardless of whether the regular-grid method or the line-integral method is used and what interpolation schemes are used, the task at hand is often to derive data at grids where no measurements or only aliased measurements are available. Given the above discussions, one has to realize the limit that the accuracy of objective analysis can reach. After elaborate and in-depth discussion of the analysis of the GATE data, Ooyama (1987) concluded his paper with the statement that "to make gold, one must start with gold."

3. A scheme used to analyze the ARM IOP data

The ARM program has two scientific objectives: (1) to improve the treatment of radiation processes in GCMs, and (2) to improve cloud parameterizations in these models (Stokes and Schwartz 1994). It is the second objective that requires the accurate knowledge of atmospheric dynamic structures at the ARM experimental testbed, including the vertical velocity and advective tendencies. One of the ARM initiatives is single-column modeling (SCM) (Randall et al. 1996; Ghan et al. 1999) which requires observational input of vertical velocity and advective tendencies of temperature and water vapor. The ARM Southern Great Plains (SGP) site has been in operation for several years, and an unprecedented amount of measurements has been collected. This section outlines an analysis scheme designed for ARM, and discusses the implementation of the scheme for the analysis of the July 1995 ARM IOP data.

a. General strategy

At the ARM SGP (Figure 4a), the number of balloon stations alone, considering the possible missing or bad soundings, is not well suited for the use of the regular grid method. If one sounding is missing, say at station B1 or B4, the whole mission is seriously compromised. Therefore, wind measurements from wind profilers are important to complement the balloon soundings. There are many NOAA wind profiler stations taking hourly winds near the SGP site. Figure 4b shows the locations of the nearest 7 NOAA profilers. It is seen that all five sounding stations actually overlap with the wind profilers. The objective analysis domain is better defined when the two additional NOAA profiler stations are used, which make the analysis domain more symmetric.

In our analysis, we use a hybrid method of combining the strengths of the regular grid method and the line integral method. The procedure is equivalent to the regular-grid method in that we establish the analysis grids based on our purpose, and we interpolate observations onto these grids. It differs from the traditional regular grid method in that we select the analysis grids to be as close as possible to the original observational stations, so that an analysis grid is never far away from a measurement station. Since most analysis grids overlap with measurement stations, actual measurements at these grids, if available, will dominate the analysis. The final analysis is therefore somewhat transparent to the original input data. A line-integral method is used in the end to derive the fluxes out of or into the analysis domain.

We further take the liberty of placing one more grids on the sides of the analysis domain in Figure 4b, as seen in Figure 4c. These auxiliary grids are added to improve the linear assumption on the line segment at the boundary, making use of measurements that are not located at the domain boundary. This method can also account for the drifting and the time delay of balloons at different stations, since these factors can be included in the interpolation schemes. A fixed pressure coordinate with vertical resolution of 50 mb is used in the current study.

Collocated with the upper air balloon and profiler measurements, there is a dense surface measurement network at the ARM SGP site, along with satellite measurements from GOES. Figure 5a shows a subset of ARM measurement platforms whose data will be used in our objective analysis. These platforms include: (i) the SMOS stations (surface meteorological observational stations) that measure precipitation, surface pressure, winds, temperature, and relative humidity. (ii) the EBBR stations (the Energy Budget Bowen Ratio) that give surface broadband net radiative flux and surface sensible and latent heat fluxes. (iii) the Oklahoma and Kansas mesonet stations measuring surface precipitation, pressure, winds, and temperature. (iv) the MWR (microwave radiometer stations) stations, which overlap with the sounding stations to measure precipitable water and total cloud liquid water. Satellite measurements of clouds and broadband radiative fluxes are available from NASA Langley at half hour frequency (Minnis et al. 1995). The $0.5°$ $\times 0.5°$ satellite grids are shown in Figure 5b. In addition, operational analysis from the NOAA mesoscale model RUC (Rapid Update Cycle) is available to us, with its grids shown in Figure 4d. Additional measurement platforms that are relevant to our objective analysis may be available for other IOPs. For example, for the 1997 Summer ARM IOP, ECOR (Eddy CORrelation) measurements of surface latent and sensible heat fluxes are available from the ECOR stations; and radar rainfall of 4km x 4km resolution covering the whole analysis domain is also available. All these measurement platforms provide a three-dimensional structure of an atmospheric column.

The upper air analysis, once performed on the twelve analysis grids in Figure 4, can be directly used to calculate the divergence and convergence of the fields. On top of this analysis, we further constrain the final upper air analysis to be consistent with surface and satellite measurements. This is achieved by making minimum possible adjustments to the initial upper air analysis at each grid so that the final analysis yields a column-integrated budgets of atmospheric mass, energy, water vapor, and momentum that are consistent with budgets derived from surface and TOA measurements. Therefore, in the final analysis, whatever comes horizontally into the atmospheric column equals to whatever comes out from the bottom and the TOA plus the increase of local storage of mass, energy, water vapor, and momentum. As described in the previous sections,

11

there is a limit in the accuracy of the initial upper air analysis. Adjustments of this initial analysis are justified as long as their magnitudes are within the expected error bounds.

Technical details of this constrained variational step has been described in Zhang and Lin (1997). Zhang and Lin (1997) only used four balloons from the October 1994 ARM IOP data to illustrate the method. When applied to the present ARM data streams as described in Figures 4 and 5, the procedure is briefly described as follow: The final analysis of atmospheric winds, temperature, and humidity at the twelve analysis grids highlighted in Figures 4 and 5, is required to satisfy the following vertical-integrated constraints:

$$< \nabla \cdot \vec{V} > = -\frac{1}{gp_s}\frac{dp_s}{dt} \tag{9}$$

$$\frac{\partial <q>}{\partial t} + < \nabla \cdot \vec{V}q > = E_s - Prec - \frac{\partial <q_l>}{\partial t} \tag{10}$$

$$\frac{\partial <s>}{\partial t} + < \nabla \cdot \vec{V}s > = R_{TOA} - R_{SRF} + LPrec + SH + \frac{\partial <q_l>}{\partial t} \tag{11}$$

$$\frac{\partial <\vec{V}>}{\partial t} + < \nabla \cdot \vec{V}\vec{V} > -f\vec{k} \times < \vec{V} > -\nabla < \phi > = \vec{\tau}_s \tag{12}$$

where the bracket represents vertical integration, $\vec{v}$ is the wind, q is the water vapor mixing ratio, $s=CpT+gz$ is the dry static energy, and p_s is the surface pressure. R is the net downward radiative flux at the TOA and at the surface (SRF). $\vec{\tau}_s$ is the surface wind stress, *Prec* is the precipitation, *SH* is the sensible heat flux, E_s is the surface evaporation, and q_l is the cloud liquid water content. All spatial derivative terms in (9)-(12) are calculated using the line-integral method. The final analysis product is derived by minimizing the following cost function:

$$I(t) = (u*-u_o)^T Q_u(u*-u_o) + (v*-v_o)^T Q_v(v*-v_o) + (q*-q_o)^T Q_q(q*-q_o) \tag{13}$$

with (9)-(12) as strong constraints. Subscript "o" in (13) denotes initial analysis from the interpolation schemes; α s are the weighting functions that need to be prescribed according to error estimates in the initial analysis.

The main conclusion from Zhang and Lin (1997) was that the magnitude of required adjustment to the initial upper air data is small, well within the expected data errors. The adjustment does not appreciably change the atmospheric state variables at each analysis grid. Yet, this adjustment makes a large impact on the derivative variables such as divergence (thus vertical velocity) and advective tendencies. This is understandable because of the relatively small size of the field

experimental domain. Error of 1 ms^{-1} on the top of measurement value 30 ms^{-1} is only about 3%. Yet, if it is divided by a horizontal distance of 100 km, the resulting error in the divergence field is 1.0e-5 s^{-1}, which could be 100% of the divergence in a typical synoptic weather system. Zhang and Lin (1997) also presented results of the impact of using separated constraints in (9)-(12) on the derivative fields. Significant impact was shown to come from the moisture and energy constraints.

Our approach therefore requires the use of many surface measurements, in particular, area averaged surface precipitation. As highlighted in Figures 4 and 5, the ARM SGP testbed provides the ideal observational network to perform this type of analysis. As will be shown later, these constraints help to reduce the sensitivity of the analysis product on the choice of interpolation schemes and their implementations.

b. The implementation for the July 1995 ARM IOP

We proceed to describe the up-to-date implementation of the above hybrid analysis scheme in ARM. From working with ARM and TOGA-COARE data, it is clear to us that some of the implementation method needs to be modified according to the actual available data streams and their quality. Therefore, certain details may change over time.

i) upper air analysis

The original upper air measurements from the five balloon stations and the 17 NOAA profilers are analyzed using the Cressman scheme with a four-dimensional length scale of [50 km, 50 km, 50 mb, 6 hours]. Iteration is carried out for three times. RUC analysis is used as our background. We will present analysis from using other analysis schemes as sensitivity tests. Since profiler winds are available only in height coordinates, atmospheric temperature and humidity (for the calculation of virtual temperature) are first analyzed using balloon soundings and RUC analysis at the profiler stations. Profiler wind measurements are then converted to pressure levels for subsequent analysis.

It is worthwhile at this point to deviate from the main course of our discussion to briefly comment on the quality of the RUC analysis for an appraisal of its relevance as a background field, and the possibility of substituting objective analysis with operational analysis. Figure 6a shows the operational analysis from the RUC model of the 500 mb northward wind component at the Central Facility. Also shown is the sounding data after a five-point running averaging. It is seen that the RUC captures the sounding data reasonably well to the first order. Figure 6b shows the difference

of the northward wind component between boundary facility B1 and the central facility C1, from both the RUC model and from the soundings. Even though the RUC model still captures the main events, the difference between the RUC model and the sounding data can be as large as fifty percent of this quantity. Therefore, the RUC analysis does not always capture the observed derivative field. Further results will be presented later in the sensitivity results.

We next highlight some heterogeneous features of the input data. In Figure 7a, the dots denote the time variation of the profiler northward wind component (the dots) at the central facility at 500 mb averaged to every 3 hours. Considerable amount of temporal variability is recorded even though the data have been averaged to three hours. Also plotted in Figure 7a is the corresponding wind measurement from the balloon sounding (the crosses). While winds from the two data sources agree with each other in capturing many broad features, they sometimes differ significantly. This is more clearly seen in the five-point running averages plotted in Figure 7a. The balloon soundings represent point measurements, while the profilers measure wind of an air layer of 500 m thick. The two are not expected to completely agree with each other. Nevertheless, the difference at the end of the IOP is beyond what we can reconcile. This observation, together with the subjectivity in the implementation interpolation scheme, re-enforces the notion that there is a limit in the accuracy of the upper air analysis. We will present sensitivity tests showing results by using balloon sounding data alone and profiler data alone in Section 4.

Compounding the measurement uncertainty is the inevitably large amount of missing data. Figure 8a shows the missing soundings at each time and height out of the five balloon stations in the July 1995 ARM IOP. Figure 8b shows the missing profiler measurements over the seven profilers that are closest to the ARM SGP. Over about half of the time, there is at least one missing balloon sounding, and there is at least one missing profiler at any time.

Before the initial analysis is performed, rigorous quality control of the input data is needed. We implemented several procedures to automatically screen both the sounding data and the profiler data. In the end, we caution that visual check of all input data is still necessary. Figures 9a and 9c show two examples of the distribution of sounding wind at boundary facility B6. The winds at one of the sounding time in each figure are suspicious only when they are visually inspected (day 181.6, 14:30 GMT, June 30, 1997, and day 193.0, 23:30 GMT, July 11, 1997). This is confirmed by the wind distribution at the nearby central facility as shown in Figures 9b and 9d. These data need to be rejected in the analysis. An example for the profiler data is shown in Figure 10, in which the time series of the two wind components are plotted at an altitude of 8 km. Aside from the large scatters, one can see a line of near-zero values in the u component. When the v

values that correspond to these near-zero u values are highlighted in Figure 10b, it is seen that most of these entries also have near-zero v component. These values stand out only in the visual check, and they need to be deleted from the data.

Considering the temporal and spatial variability of the profiler winds that are less coherent than those from the sounding data, we used a half weight for the profiler data and full weight for the sounding data in the Cressman procedure. Thus, the analysis at C1, say, takes the largest weight from the sounding data at C1, and half of that weight from the profiler at C1, and smaller weights from other pressure levels, time steps, and other stations. An example of the finally analyzed wind at 500 mb at C1 is given in Figure 7b, along with the five-point running averaged sounding data, profiler data at the same station. It is seen that this analysis primarily follows the sounding data, and it is smoother than the sounding data. As can be imagined, there is some subjectivity in giving more or less weight to the sounding data than to the profiler data. The initial analysis of the derivative fields is somewhat sensitive to this subjective choice, which will be reduced by the variational constraining process.

ii) constraint variables

All terms on the right hand side of equations (9)-(12) should be the area-averaged quantities within the analysis domain, since the advective transport terms on the left sides only describe area-averaged quantities. Sampling of surface measurements in calculating area-averaged quantities then becomes a concern, particularly for variables that have large spatial variabilities such as precipitation. Figure 11a shows the precipitation averaged over 8 SMOS stations within the analysis domain, and that averaged over 77 Oklahoma and Kansas mesonet stations within the same domain. It is seen that they could differ by a factor of two, with rainfall at the SMOS stations systematically larger than those from the mesonet stations. A large sample of stations is needed to define the area averaged precipitation.

To avoid biases of using overcrowding measurement stations in some areas, we lay the 0.5°x0.5° GOES grids in Figure 5b over the analysis domain in Figure 5a, and then derive the required quantities in each small grid box. If there are actual measurements within a 0.5°x0.5° grid box, simple arithmetic averaging is used to obtain the sub-grid means. Some variables are available from several instruments. For example, surface winds are available from the sounding stations, the SMOS stations, the Oklahoma and Kansas mesonet stations. They are merged in the arithmetic averaging process. If there is no actual measurement in a 0.5°x0.5° box, then the Barnes scheme is used with the length scale of (Lx=50 km, Ly=50 km, Lt=6 hours) to fill the missing data.

Missing broadband shortwave measurement is filled by applying this procedure to the albedo and then converting it to radiative flux. Domain averages of these quantities are obtained by using values from the 0.5°x0.5° grid boxes within the analysis domain.

Figure 11b shows the derived area-averaged precipitation when all measurement stations are used in the calculation of gridded values. Also plotted are values when only half, and one third of the measurement stations are used. It is seen that the area-averaged precipitation asymptotes as the number of surface station increases to approach 50%, which is over 60 stations.

All other surface variables are calculated using the same procedure. Figure 12a shows the surface latent and sensible heat fluxes from the EBBR stations. Figure 12b shows the atmospheric net radiative heating. Together with the precipitation, these measurements form the constraints on the horizontal transport and time variability of heat and water vapor within the atmospheric column.

iii) error variances

The constrained variational step requires the specification of error estimates to determine the weights in the cost function (13). The relative magnitudes of these estimates, rather than their absolute magnitudes, enter into the minimization procedure. Zhang and Lin (1997) specified these weights according to instrument and measurement uncertainty estimates provided by the ARM instrument team. These weights were also used in the analysis used for the first SCM intercomparison study (Ghan et al. 1999). This specification has been recently modified to include a component that is related with the observed variance of atmospheric state variables in the sounding data. The standard deviations of the state variables are shown in Figure 13. Variability of the winds generally increases with height, with the u component having larger variability than the v component. These standard deviations are multiplied by a factor of 20% to account for aliasing errors; they are then added to the instrument and measurement estimates used in Zhang and Lin (1997), which are 0.5 ms^{-1} for winds, 0.2 K for temperature, and 3% of the specific humidity. As one can see, there is some degree of subjectivity in the selection of the error estimate. Sensitivity tests of the analysis will be presented.

4. Analysis of the July 1995 ARM IOP data and its sensitivity

The scheme described in the previous section, together with the surface and TOA measurements, is used to calculate the derivative fields for the July 1995 ARM IOP. This is the IOP for which a single-column modeling intercomparison has been carried out (Ghan et al. 1999). The IOP starts

from July 18, 1995 to August 4, 1995. The large-scale circulation over Northern America can be divided into three periods. In the beginning period from July 17-July 28, a broad trough extending from northern Canada covers most of the North America continent. The upper air circulation over the most of the United States is characterized by westerlies accompanying the bottom of the trough. The ARM SGP site is at the southern edge of this trough, and it is near the northern edge of the subtropical high as denoted by the 588 line on the 500 mb map. Weather condition at the ARM site is therefore under influences of both the subtropical high and the middle latitude circulations. It is under this large-scale circulation that several strong precipitation events, as seen from the precipitation in Figure 11, took place.

The above circulation pattern remained stable for over a week, and then it was disrupted by a developing ridge in the North Pacific that was further driven by a developing upstream trough starting from July 26, 1995 (day 207). This ridge then moved over the North America continent. Upper level winds over the ARM SGP are southeastward. This period corresponds to the relatively quite period in Figure 11.

The clear-sky period ended when the ridge moved out and the previously mentioned developing trough moved into the continent. Weather at the SGP is again under the control of a trough. This trough-controlled condition differed from the first period in that the bottom of the trough coincided with a tropical depression. Possibly because of this coincidence, the depression developed into a tropical hurricane. The ARM SGP is therefore located at the intersection of the middle latitude trough and the tropical hurricane. Strong and continuous precipitation was observed (Figure 11).

The mesoscale circulation over the ARM SGP is featured by the strong diurnal variation of lower level winds that are confined below 800 mb. The lower level jet often develops at the local evening, with wind blowing from southwest to northeast in the night. The wind direction turns gradually to align with the westerlies during the day time. Associated with this change in the wind pattern, there is a strong diurnal variation of water vapor and temperature advections. The diurnal variation of the jet appears to occur in a scale larger than the convective clusters. The continental scale land-sea breeze is thus possibly responsible to the formation of the diurnal winds in this region.

Figures 14a and 14b show the time-pressure cross sections of the analyzed temperature and water vapor mixing ratio. It is seen that there are strong diurnal variations of temperature and water vapor near the surface below 850 mb. There is a general warming and drying in the second period when the atmospheric ridge is over the SGP. Other than these variations, the temperature field is

relatively flat. The water vapor field is different; there are clear signatures in the upper tropospheric moisture associated with the convective events.

The analyzed wind divergence and vertical velocity are shown in Figures 14c and 14d. It is seen that episodes of strong upward motion occur in the first period, which appear to be diurnally locked. There is overall downward motion starting around July 25 (day 206). The last period corresponds to relatively uniform upward motion. The transition altitude of wind convergence to wind divergence is higher for the episodic convective events than that for the continuous precipitation period. This altitude is where the upward motion peaks.

Attribution to the large-scale advection and Q1, Q2 (Yanai et al., 1973) of the observed changes of temperature and moisture, and analysis of the contributions of latent heating and turbulent heating, will be presented in a separate study. In this paper, we focus on the robustness of the objectively analyzed product. To reduce the length of the paper, we will only employ wind divergence and vertical velocity to examine the data sensitivities. Four sets of experiments are carried. Results will be presented for analyses with mass constraint only and with all constraints of (9)-(12). The first set of experiments is designed to test the impact of the free parameters in the Barnes and Cressman schemes, namely the length scale in the weighting function and the number of iterations. As shown in Section 2, these parameters are equivalent to the smoothing and filtering in the statistical interpolation scheme. The second set of experiments is designed to examine the sensitivity of the analysis to the use of the Barnes scheme, the Cressman scheme, and the statistical interpolation scheme. The third set is to test the sensitivity of the analysis on the availability of input data, including profiler data, sounding data, and RUC output. The fourth set tests the sensitivity of the data on the constraint variables in (9)-(12) and on the weighting coefficients in the cost function (13).

a. Length scale and number of iterations

For brevity, we only present results by using the analyses from the Barnes scheme. The vertical velocities for the July 1995 ARM IOP from the Barnes scheme, analyzed with mass constraint only, are shown in Figures 15a to c. In Figure 15a, only one iteration is performed in the Barnes scheme with a four-dimensional length scale of (Lx=50 km, Ly=50 km, Lp=50 mb, Lt=3 hr). Figure 15b is the same except that a larger length scale of (Lx=100 km, Ly=100 km, Lp=50 mb, Lt=6 hours) is used. Figure 15c is the same and Figure 15b except with three iterations. Consistent with what can be expected from Figure 3a-c, much smoother and weaker vertical velocity is produced with a larger length scale, and with less iteration. It is noted that the

difference in the vertical velocity is significant between the analyses (e.g., the intensity of the continuous upward motion after day 212).

Figures 15d-15f show analyses of vertical velocity corresponding to the same three experiments except with all constraints of (9)-(12) imposed. It is seen that use of the surface and TOA measurements significantly reduces the sensitivity of the analyses. We note, however, that differences still exist between the analyses, which could be large for some applications. For example, the upward motion in day 203 in Figure 15e is much weaker than that in Figure 15d. We also note that the analyzed vertical velocity from using all constraints is quite different from that without using these constraints.

b. Interpolation schemes

We next examine the sensitivity of the analyzed data to the use of the Barnes scheme, the Cressman scheme, and the statistical interpolation scheme. As shown previously in Figure 3, with each analysis scheme, we carried three experiments, using different length scales and number of iterations for the Barnes scheme and the Cressman scheme, and using different de-aliasing method for the statistical method. We then averaged results from these three experiments for each scheme. Figures 16a-16c show the analysis of vertical velocity from each scheme with constraint of mass conservation only. While all objective schemes capture the general features of the events, there exist large differences. For example, the upward motion in day 213 from the Barnes scheme is weaker than that in the Cressman scheme, which is further weaker than that in the statistically interpolated product. Furthermore, an additional upward motion event is analyzed at day 202 in the statistical product, and more small scale features are present in this product. When the additional constraints of (9)-(12) are imposed on the analysis, the sensitivities of the analyses on the schemes are reduced (Figures 16d-16f), even though the constraining step cannot eliminate all the differences. This is seen in Figure 16f in which more structures are still seen than in the other two analyses.

c. Data sources

We now experiment with using only sounding data, only profiler data, and the RUC model output as our upper air data sources. Figures 17a-17c show the analysis with mass constraint only from using: (a) sounding data only, (b) profiler data only, and (c) RUC model output. Since the profiler data are only available for winds, temperature and water vapor analyses in Figure 17b are taken from RUC. It is seen that the analysis from using the profiler data alone is quite different.

The vertical velocity is too weaker and, consistent with Figure 7a, the analysis from profiler data even missed the major event in the last period.

Figures 18a-18c show the distribution of the wind divergence corresponding to Figures 17a-17c. The profiler winds badly missed the mid-tropospheric wind convergence in the last period. It is also seen that the magnitude of the wind divergence in the RUC model is generally weaker than that in the sounding data, consistent with the snapshot shown in Figure 6.

Figures 17d-f show the same analysis of vertical velocity when all constraints are used. Consistent with the previous sensitivity results, the differences from using the different data sources are reduced. Substantial difference, however, still exists. For example, the upward motion from using the profiler data alone appears to be at a lower altitude than that in the sounding data. These differences are more clearly seen in the wind divergence field, which are shown in Figures 18d-f. The variational procedure has adjusted the profiler winds to force an upward motion. Yet, the largest adjustments are made near the surface where adjustment to the water vapor balance is most efficient. Obvious difference also exists for the RUC model, in which the area of convergence in the last period extends into a deeper layer. Thus, the variational constraints can reduce the sensitivity of the data product, but it can not substitute for the quality of the original upper air data.

d. Constraint variables and error variances

We now examine the sensitivity of the analysis to the constraining parameters of the variational procedure. We present results from three experiments. In the first experiment, we changed the area-averaged precipitation by using estimates from one third of the observational stations. The difference in precipitation has been shown in Figure 11c. We note that characterizing the error uncertainty in the actual precipitation warrants separate study, and here we simply consider the precipitation difference in Figure 11c as a rough estimate of the uncertainty. In the second experiment, we used the Sib (Simple Biosphere Model) derived surface latent and sensible heat fluxes as a sensitivity test. These fluxes are plotted in Figure 12c, they differ significantly from those of the EBBR measurements. Again the difference is considered as an error estimate in the surface turbulent fluxes. In the third experiment, we change the weights in the cost function of (13) to those used in Zhang and Lin (1997), in which the weights become invariant with height for winds and temperature. In the control case, as described in section 3c, the weights were prescribed based on measurement uncertainties and small-scale aliasing calculated from the standard deviations shown in Figure 13.

Figure 19a shows the analysis of vertical velocity when the precipitation is derived from one third of the observational stations. This distribution of the vertical velocity is very similar to that in the control case in which all precipitation data are used. When we look at the difference field, however, Figure 19d shows that the intensity of precipitation has a direct impact on the magnitude of the vertical velocity. Stronger precipitation corresponds to stronger upward motion. This experiment suggests that our analyzed data at the ARM SGP are subject to uncertainty related with the use of the surface precipitation, and yet this uncertainty is likely to be smaller than those from the analysis schemes without the use of additional constraints.

Figure 19b shows the vertical velocity when the sensible and latent heat fluxes from the EBBR stations are replaced by using those from the Sib model (courtesy of Chris Doran, PNL). Figure 19e shows the difference in vertical velocity from using the different turbulent surface fluxes. Reduction of latent heat flux and increase of sensible heat flux, from using the Sib fluxes, both lead to increased upward motion, since the former would require larger convergence of moisture into the column and therefore stronger upward motion, while the latter requires more divergence of dry static energy and thus stronger divergence in the upper layers. The uncertainty in the analyzed data that is related with the turbulent surface fluxes, however, is seen to be generally smaller than those from the precipitations.

We finally show the sensitivity of the vertical velocity on the choice of the weighting coefficients in the variational cost function (13). Figures 19c and 19f (d) show the analysis of vertical velocity and its difference with the control experiment. Unlike the surface constraints, the weighting coefficients change the vertical structure of the analysis, even though the overall magnitude appears to be smaller than those from other perturbations.

5. Summary and Conclusions

We have discussed the strength and weakness of the regular-grid method, the line-integral method, and the following interpolation schemes: the Barnes scheme, the Cressman scheme, the statistical interpolation scheme, for the derivation of the atmospheric wind divergence, vertical velocity, and advective tendencies from field experimental data. Considerable differences in the analyzed products can be attributed to the selection of analysis schemes, and even more to the actual implementation of a particular scheme.

We have introduced a hybrid approach to combine the strength of the regular grid method and the line-integral method for the initial analysis of ARM data. This combination enables better sampling

of the data than the simple line-integral method, and yet allows more transparent connections between the analysis and the actual measurements than in the traditional regular grid method. The initial analysis is then subject to a variational constraining process that ensures the conservation of the column integrated mass, energy and water vapor budgets in the analysis product. In addition to the use of upper air data, surface measurements of precipitation, sensible and latent heat fluxes are all used in the objective analysis.

Four sets of experiments are carried out to examine the sensitivity of the analyzed product to the analysis methods and data sources, with and without the column-integrated constraints of water and energy budgets. It is shown that the constraining process significantly reduces the sensitivity of the final data products. It is also shown that uncertainty of the vertical velocity field caused by uncertainties in the variational procedure is smaller than those from using different schemes, or from using different data sources without the surface and TOA measurements. In the constrained variational analysis, accuracy of area-averaged precipitation appears to have the largest impact on the data products. Improvements on the surface analysis, therefore, become an integral part in the ARM objective analysis of upper air measurements.

REFERENCES

Barnes, S. L., 1964: A technique for maximizing details in numerical map analysis, *J. Appl. Meteor.*, **3**, 396-409.

Betts, A. K., 1973: A composite mesoscale cumulonimbus budget, *J. Atmos. Sci.*, **30**, 597-610.

Bradley, F. and R. Weller, Eds., 1995: Joint workshop of the TOGA COARE flux and atmospheric working groups, 11-13 July 1995. (available from TOGA COARE International Project Office).

Chen, Y.-L., and J. Lin, 1995: Large-scale conditions favorable for the development of heavy rainfall during TAMEX IOP 3, *Mon. Wea. Rev.*, **123**, 2978-3002.

Cotton, W. R., M.-S. Lin, R. L. McAnelly, and C. J. Tremback, 1989: A composite model of mesoscale convective complexes, *Mon. Wea. Rev.*, **117**, 765-783.

Cressman, G. P., 1959: An operational objective analysis scheme, *Mon. Wea. Rev.*, **87**, 367-374.

Daley, R., 1991: Atmospheric Data Analysis, Cambridge U. Press, 457 pp.

Davies-Jones, R. P., 1993: Useful formulas for computing divergence, vorticity, and their errors from three or more stations, *Mon. Wea. Rev.*, **121**, 713-725.

Esbensen, S. K., 1975: A analysis of subcloud-layer heat and moisture budgets in the Western Atlantic Trades, *J. Atmos. Sci.*, **31**, 1921-1933.

Esbensen, S. K., E. I. Tollerud, and J.-H. Chu, 1982: Cloud-cluster-scale circulations and the vorticity budget of synoptic-scale waves over the eastern Atlantic intertropical convergence zone, *Mon. Wea. Rev.*, **110**, 1677-1692.

Frank, W. M., 1979: Individual time period analysis over the GATE ship array, *Mon. Wea. Rev.*, **107**, 1600-1616.

Gallus, W. A., and R. H. Johnson, 1991: Heat and moisture budgets of an intense midlatitude squall line, *J. Atmos. Sci.*, **48**, 122-146.

Gallus, W. A., and R. H. Johnson, 1992: The momentum budget of an intense midlatitude squall line, *J. Atmos. Sci.*, **49**, 422-450.

Ghan, S. et al., 1999: An intercomparison of single column model simulations of summertime midlatitude continental convection. Submitted to *J. Geophys. Res.*

He, H. Y., J. W. McGinnis, Z. S. Song, and M. Yanai, 1987: Onset of the Asian summer monsoon in 1979 and the effect of the Tibetan Plateau, *Mon. Wea. Rev.*, **115**, 1966-1995.

Johnson, R. H., 1976: The role of convective-scale precipitation downdrafts in cumulus and synoptic-scale interactions, *J. Atmos. Sci.*, **33**, 1890-1910.

Johnson, R. H., and G. S. Young, 1983: Heat and moisture budgets of tropical mesoscale anvil clouds, *J. Atmos. Sci.*, **40**, 2138-2147.

Johnson, R. H., and J. F. Bresch, 1991: Diagnosed characteristics of precipitation systems over Taiwan during the May-June 1987 TAMEX, *Mon. Wea. Rev.*, **119**, 2540-2557.

Jorgensen, D. P., M. A. LeMone, and J.-D. Jou, 1991: Precipitation and kinematic structure of an oceanic mesoscale convective system. Part I: Momentum transport and generation, *Mon. Wea. Rev.*, **119**, 2638-2653.

Kuo, Y.-H., and R. A. Anthes, 1984a: Accuracy of diagnostic heat and moisture budgets using SESAME-79 field data as revealed by observing system simulation experiments, *Mon. Wea. Rev.*, **112**, 1465-1481.

Kuo, Y.-H., and R. A. Anthes, 1984bMesoscale budgets of heat and moisture in a convective system over the central United States, *Mon. Wea. Rev.*, **112**, 1482-1497.

Kuo, Y.-H., L.-S. Cheng, and R. A. Anthes, 1985: Mesoscale analysis of the Sichuan flood catastrophe, 11-15 July 1981, *Mon. Wea. Rev.*, **114**, 1984-2003.

Lafore, J.-P., J.-L. Redelsperger, and G. Jaubert, 1988: Comparison between a three-dimension simulation and doppler radar data of a tropical squall line: transports of mass, momentum, heat, and moisture, *J. Atmos. Sci.*, **45**, 3483-3500.

Leach, M. J., J. Yio, and R. T. Cederwall, 1996: Estimation of errors in objectively analyzed fields and sensitivity to number and spacing of stations. Proceedings of the sixth Annual ARM Science Meeting, San Antonio, TX., DOE CONF-9603149, 149-151.

Lee, C. S., 1989: Observational analysis of tropical cyclogenesis in the western North Pacific. Part II: Budget analysis, *J. Atmos. Sci.*, **46**, 2599-2616.

LeMone, M. A., 1983: Momentum transport by a line of cumulonimbus, *J. Atmos. Sci.*, **40**, 1815-1834.

LeMone, M. A., G. B. Barnes, and E. J. Zipser, 1984: Momentum flux by lines of cumulonimbus over the tropical oceans, *J. Atmos. Sci.*, **41**, 1914-1932.

LeMone, M. A., and M. W. Moncrieff, 1994: Momentum and mass transport by convective bands: comparisons of highly idealized dynamical models to observations, *J. Atmos. Sci.*, **51**, 281-305.

Lewis, J. M., 1975: Test of the Ogura-Cho model on a prefrontal squall line case, *Mon. Wea. Rev.*, **103**, 764-778.

Lin, X., and R. H. Johnson, 1994: Heat and moisture budgets and circulation characteristics of a frontal squall line, *J. Atmos. Sci.*, **51**, 1661-1681.

Lin, X., and R. H. Johnson, 1996a: Kinematic and thermodynamic characteristics of the flow over the Western Pacific warm pool during TOGA COARE, *J. Atmos. Sci.*, **53**, 695-715.

Lin, X., and R. H. Johnson, 1996b: Heating, moistening, and rainfall over the western Pacific warm pool during TOGA-COARE, *J. Atmos. Sci.*, **53**, 3367-3383.

Lin, Y.-J., T.-C. C. Wang, R. W. Pasken, H. Shen, and Z.-S. Deng, 1990: Characteristics of a subtropical squall line determined from TAMEX dual-doppler data. Part II: Dynamic and thermodynamic structures and momentum budgets, *J. Atmos. Sci.*, **47**, 2382-2399.

Lord, S. J., and J. L. Franklin, 1990: The environment of Hurricane Debby (1982). Part II: Thermodynamic fields, *Mon. Wea. Rev.*, **118**, 1444-1459.

Luo, H. B., and M. Yanai, 1984: The large-scale circulation and heat sources over the Tibetan Plateau and surrounding areas during the early summer of 1979. Part II: Heat and moisture budgets, *Mon. Wea. Rev.*, **112**, 966-989.

McBride, J. L., B. W. Gunn, G. J. Holland, T. D. Keenan, and N. E. Davidson, 1989: Time series of total heating and moistening over the Gulf of Carpentaria radiosonde array during AMEX, *Mon. Wea. Rev.*, **117**, 2701-2713.

McNab, A. L., and A. K. Betts, 1978: A mesoscale budget study of cumulus convection, *Mon. Wea. Rev.*, **106**, 1317-1331.

Miller, B. L., and D. G. Vincent, 1987: Convective heating and precipitation estimates for the tropical South Pacific during FGGE, 10-18 January 1979, *Q. J. R. Meteorol. Soc.*, **113**, 189-212.

Minnis, P., W. L. Smith, D. P. Garber, J. K. Ayers, and D. R. Doeling, 1995: Cloud properties derived from GOES-7 for spring 1994 ARM Intensive Observing Period using version 1.0.0 of ARM satellite data analysis program, NASA Reference Publication 1366.

Nitta, T., 1972: Energy budget of wave disturbances over the Marshall Islands during the years of 1956 and 1958, *J. Meteor. Soc. Japan*, **50**, 71-84.

Nitta, T., 1977: Response of cumulus updraft and downdraft to GATE A/B-scale motion systems, *J. Atmos. Sci.*, **34**, 1163-1186.

Nitta, T., and S. Esbensen, 1974: Heat and moisture budget analyses using BOMEX data, *Mon. Wea. Rev.*, **102**, 17-28.

Ooyama, K., 1987: Scale-controlled objective analysis, *Mon. Wea. Rev.*, **115**, 2476-2506.

Pedigo, C. B., and D. G. Vincent, 1990: Tropical precipitation rates during SOP-1, FGGE, estimated from heat and moisture budgets, *Mon. Wea. Rev.*, **118**, 542-557.

Randall, D. A., K. M. Xu, R. C. J. Somerville, and S. Iacobellis, 1996: Single-column models and cloud ensemble models as links between observations and climate models, *J. Climate*, **9**, 1683-1697.

Redelsperger, J. L., et al., 1999: A GCSS model intercomparison for a tropical squall line observed during TOGA-COARE, Part I: cloud resolving models, Submitted to *Quart. J. Roy. Meteor. Soc.*

Reed, R. J., and E. E. Recker, 1971: Structure and properties of synoptic-scale wave disturbances in the equatorial Western Pacific, *J. Atmos. Sci.*, **28**, 1117-1133.

Sanders, F., and K. A. Emanuel, 1977: The momentum budget and temporal evolution of a mesoscale convective system, *J. Atmos. Sci.*, **34**, 322-330.

Song, J.-L., and W. M. Frank, 1983: Relatinoships between deep convection and large-scale processing during GATE, *Mon. Wea. Rev.*, **111**, 2145-2160.

Stokes, G. M., and S. E. Schwartz, 1994: The Atmospheric Radiation Measurement (ARM) Program: programmaric background and design of the Cloud And Radiation Test Bed, *Bull. Amer. Meteor. Soc.*, **75**, 1201-1221.

Sui, C.-H., and M. Yanai, 1986: Cumulus ensemble effects on the large-scale vorticity and momentum fields of GATE. Part I: Observational evidence, *J. Atmos. Sci.*, **43**, 1618-1642.

Thompson, R. M., S. W. Payne, E. E. Recker, and R. J. Reed, 1979: Structure and properties of synoptic-scale wave disturbances in the intertropical convergence zone of the eastern Atlantic, *J. Atmos. Sci.*, **36**, 53-72.

Webster, P. J., and R. Lukas, 1992: TOGA COARE: The coupled ocean-atmosphere response experiment. Bull. Amer. Meteor. Soc., **73**, 1377-1416.

Wu, X., 1993: Effects of cumulus esemble and mesoscale stratiform clouds in midlatitude convective systems, *J. Atmos. Sci.*, **50**, 2496-2518.

Wu, X., and M. Yanai, 1994: Effects of vertical wind shear on the cumulus transport of momentum: observations and parameterization, *J. Atmos. Sci.*, **51**, 1640-1660.

Yanai, M., S. K. Esbensen, and J. H. Chu, 1973: Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmos. Sci.*, **30**, 611-627.

Yanai, M., and Li., 1994: Mechanism of heating and the boundary layer over the Tibetan Plateau. *Mon. Wea. Rev.*, **122**, 305-323.

Zhang, M. H., and J. L. Lin, 1997: Constrained variational analysis of sounding data based on column-integrated conservations of mass, moisture, energy and momentum: approach and application to ARM measurements, *J. Atmos. Sci.*, **54**, 1503-1524.

Figure Captions

Figure 1 : Schematic figure of the regular grid method versus line-integral method.

Figure 2: Drifting location of balloons, relative to the ground stations, launched at the four ARM boundary facilities (B1, B4, B5, B6) at 11:30 GMT, July 22, 1995

Figure 3: Analysis of the eastward wind component (ms^{-1}). (a)-(c): Barnes scheme. (a) L=50 km, one iteration. (b) L=200 km, one iteration. (c). L=200 km, four iterations. (g) standard statistical method. (h) with smoothed true field. (i) with smoothed true field and filtered measurements.

Figure 4: Locations of the ARM upper air data streams and the analysis grids. (a) sounding stations. (b) seven profiler stations. (c) Eleven profiler stations and the twelve analysis grids in the hybrid approach. (c) RUC grids overlaid on other grids.

Figure 5: (a) ARM surface data streams. (b) GOES grids over the analysis domain.

Figure 6: Northward wind component from sounding data and RUC analysis at 500 mb (ms^{-1}). (a) sounding data versus RUC analysis at the central facility C1. (c) wind difference between two stations C1 and B1 in the sounding data and in the RUC analysis.

Figure 7: (a) Sounding and profiler wind component v at 500 mb at C1 in July 1995. (b) The analyzed wind versus measurements at C1. Unit: ms^{-1}.

Figure 8: Missing data. (a) out of five soundings. (b) out of seven profilers.

Figure 9: Sounding wind at B6 and C1 (ms^{-1}). (a) wind at B1 on June 30, 1997. (b) wind at C1 on June 30, 1997. (c) wind at B1 on July 11 to July 12, 1997. (d) wind at C1 on July 11 to July 12, 1997.

Figure 10: profiler wind components at 8 km at the central facility in July 1995 ($m\ s^{-1}$). (a) for u, and (b) for v.

Figure 11: Precipitation ($mm\ hour^{-1}$). (a) 8 SMOS stations and 77 Oklahoma and Kansas mesonet stations. (b) individual stations. (c) Averaged gridded analysis with reduced number of stations.

Figure 12: Column--integrated fluxes. (a) surface latent and sensible heat fluxes ($W\ m^{-2}$). (b) radiative heating ($W\ m^{-2}$). (c) surface latent and sensible heat fluxes from Sib model.

Figure 13: Standard deviation of atmospheric variables as function of pressure. (a) u and v (ms^{-1}). (b) temperature (K). (c) water vapor mixing ratio ($g\ kg^{-1}$).

Figure 14: Time-pressure distributions of analyzed fields. (a) temperature (K), (b) water vapor mixing ratio ($g\ kg^{-1}$), (c) wind divergence (1.0e-5 s^{-1}). (d) vertical velocity (mb $hour^{-1}$),

Figure 15: Sensitivity of the analysis of vertical velocity (mb $hour^{-1}$) to the implementation of the Barnes scheme. (a)-(c): with mass balance constraint only. (a) (Lx,Ly,Lp,Lt)=(50

km, 50 km, 50 mb, 3 hours), first iteration. (b) (Lx,Ly,Lp,Lt)=(100 km, 100 km, 50 mb, 6 hours), first iteration. (c) (Lx,Ly,Lp,Lt)=(100 km, 100 km, 50 mb, 6 hours), three iterations. (d)-(f): same as (a)-(c) except with all constraints imposed.

Figure 16: Sensitivity of the analysis of vertical velocity (mb hour^{-1}) to the interpolation schemes. (a)-(c) with mass balance constraint only. (a) Barnes scheme. (b) Cressman scheme. (c) statistical interpolation scheme. (d)-(f): same as (a)-(c) except with all constraints imposed.

Figure 17: Sensitivity of the analysis of vertical velocity (mb hour^{-1}) to upper air data sources. (a)-(c): with mass balance constraint only. (a) sounding data alone. (b) profiler data alone. (c) RUC analysis. (d)-(f): same as (a)-(c) except with all constraints imposed.

Figure 18: Sensitivity of the analysis of wind divergence (1.0e-5 s^{-1}) to upper air data sources. (a)-(c): with mass balance constraint only. (a) sounding data alone. (b) profiler data alone. (c) RUC analysis. (d)-(f): same as (a)-(c) except with all constraints imposed.

Figure 19: Sensitivity of the analysis of vertical velocity (mb hour^{-1}) to parameters in the variational procedure. (a) from using precipitation with one third of the stations sampled. (b) from using the SIB latent and sensible heat fluxes. (c) from using pressure-invariant error estimates. (d)-(f): same as (a)-(c) except for the difference of vertical velocity with the control experiment.

Table 1. Previous objective analysis of large-scale vertical velocity and advective tendencies

Asian Monsoon	North America
FGGE II-b Luo & Yanai (1984)	Oklahoma squall line Lewis (1975)
FGGE II-b He et. al. (1987)	Oklahama Sanders & Emanuel (1977)
QXPMEX Yanai & Li (1994)	Florida Johnson (1976)
1981 Flood Catastrophe Kuo et. al. (1985)	NHRE - McNab & Betts (1978)
WMONEX Johnson & Young (1983)	SESAME Kuo & Anthes (1984 a,b)
TAMEX Johnson & Bresch (1991)	SESAME Wu & Yanai (1994)
TAMEX Chen & Li n(1995)	OK PRE-STORM Gallus & Johnson (1991, 1992)
TAMEX Jorgensen et. al. (1991)	OK PRE-STORM Wu (1993)
TAMEX Lin et. al. (1990)	OK PRE-STORM Lin & Johnson (1994)
	MCC composite Cotton et. al. (1989)
	CCOPE LeMone & Moncrieff (1994)
Eastern Atlantic	**Western Atlantic**
GATE III Nitta (1977)	BOMEX Nitta & Esbensen (1974)
GATE III Thompson et. al. (1979)	
GATE III Frank (1979)	Esbensen (1975)
GATE I-III Song & Frank (1983)	
GATE III Ooyama (1987)	
GATE III Esbensen et al. (1982)	
GATE LeMone (1983)	
LeMone et. al. (1984)	
GATE Sui & Yanai (1986)	
Western Pacific	**South Pacific**
Marshall Is. Reed & Recker (1971)	FGGE III-b Miller & Vincent (1987)
Marshall Is. Nitta (1972)	
Marshall Is. Yanai et. al. (1973)	
TOGA COARE Lin & Johnson (1996)	
TOGA COARE Frank et al. (1997)	
Whole tropics	**Tropical cyclones**
FGGE III-b Pedigo & Vincent (1990)	Western Pacific composite Lee (1989)
	Hurricane Debby Lord & Franklin (1990)
Australian Monsoon	**South America**
AMEX McBride et. al. (1989)	VIMHEX Betts (1973)
Africa	
COPT81 Lafore et. al. (1988)	

AMEX -- Australian Monsoon Experiment
BOMEX -- Barbados Oceanographic and Meteorological Experiment
CCOPE -- Cooperative Covective Precipitation Experiment
COPT -- Convection Profonde Tropicale
FGGE -- First GARP Global Experiment
GATE -- Global Atmosphere Research Program (GARP) Atlantic Tropical Experiment
MONEX -- Monsoon Experiment
NHRE -- National Hail Research Experiment
OK PRE-STORM -- Oklahoma-Kansas Preliminary Regional Experiment for STORM-Central
QXPMEX -- Qinghai-Xizang Plateau Meteorology Experiment
SESAME -- Severe Environmental Storm and Mesoscale Experiment
TAMEX -- Taiwan Area Mesoscale Experiment
TOGA COARE -- TOGA Coupled Ocean Atmosphere Response Experiment ((COARE)
VIHMEX -- Venezuelan International Meteorological and Hydrological Experiment

(a) Line integral method
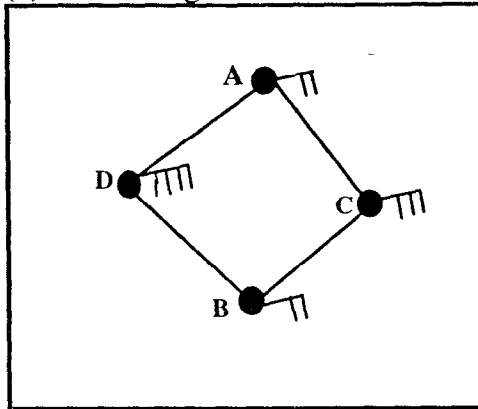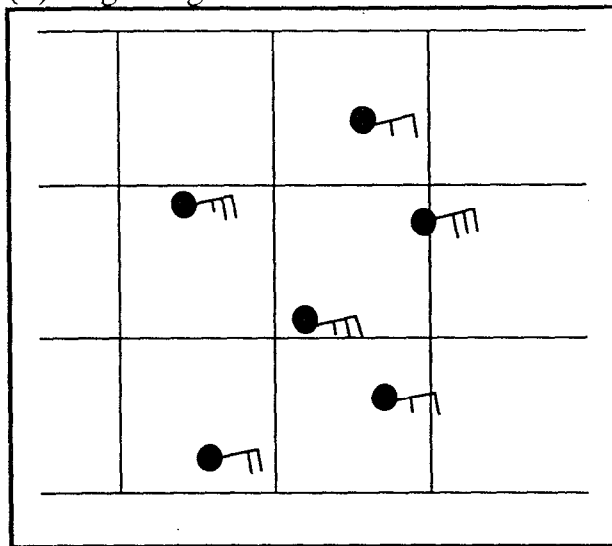

(b) Regular grid method

Figure 1. Schematic figure of the line-integral method and the regular grid method

Drifting of Balloons Launched at
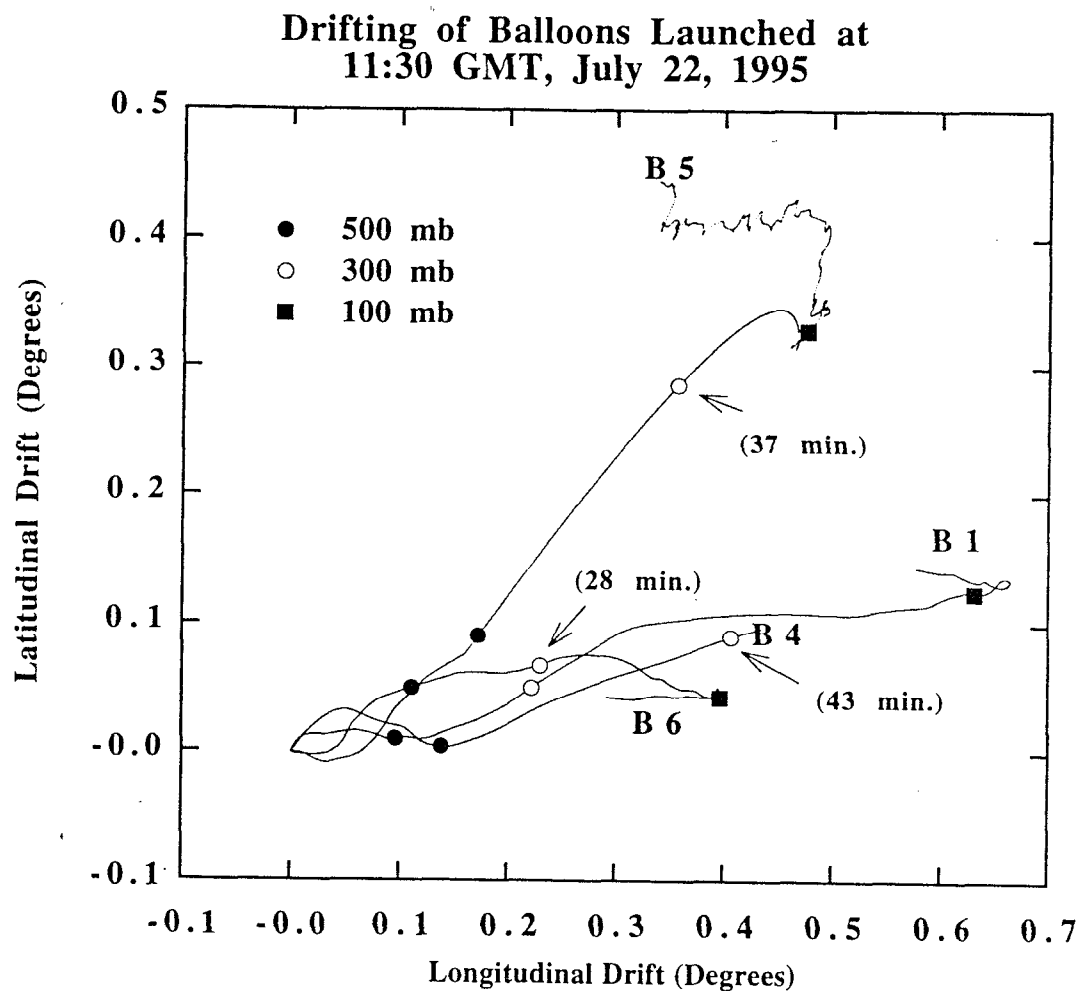## 11:30 GMT, July 22, 1995

Figure 2 : Changes of balloon locations relative to ground site at the four ARM SGP boundary facilities launched at 11:30 GMT, July 22, 1995. Also plotted are locations of balloons at three pressure levels.
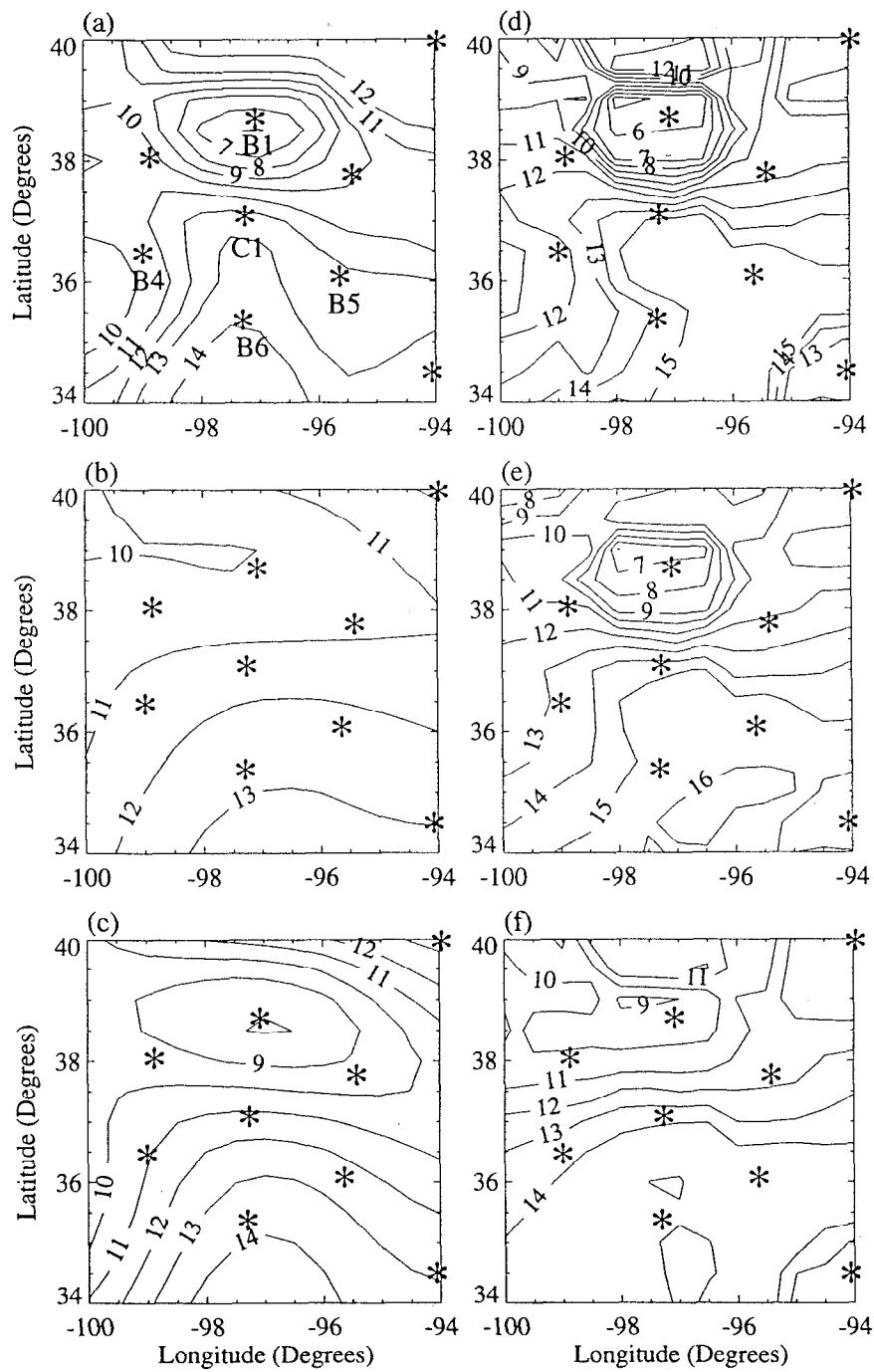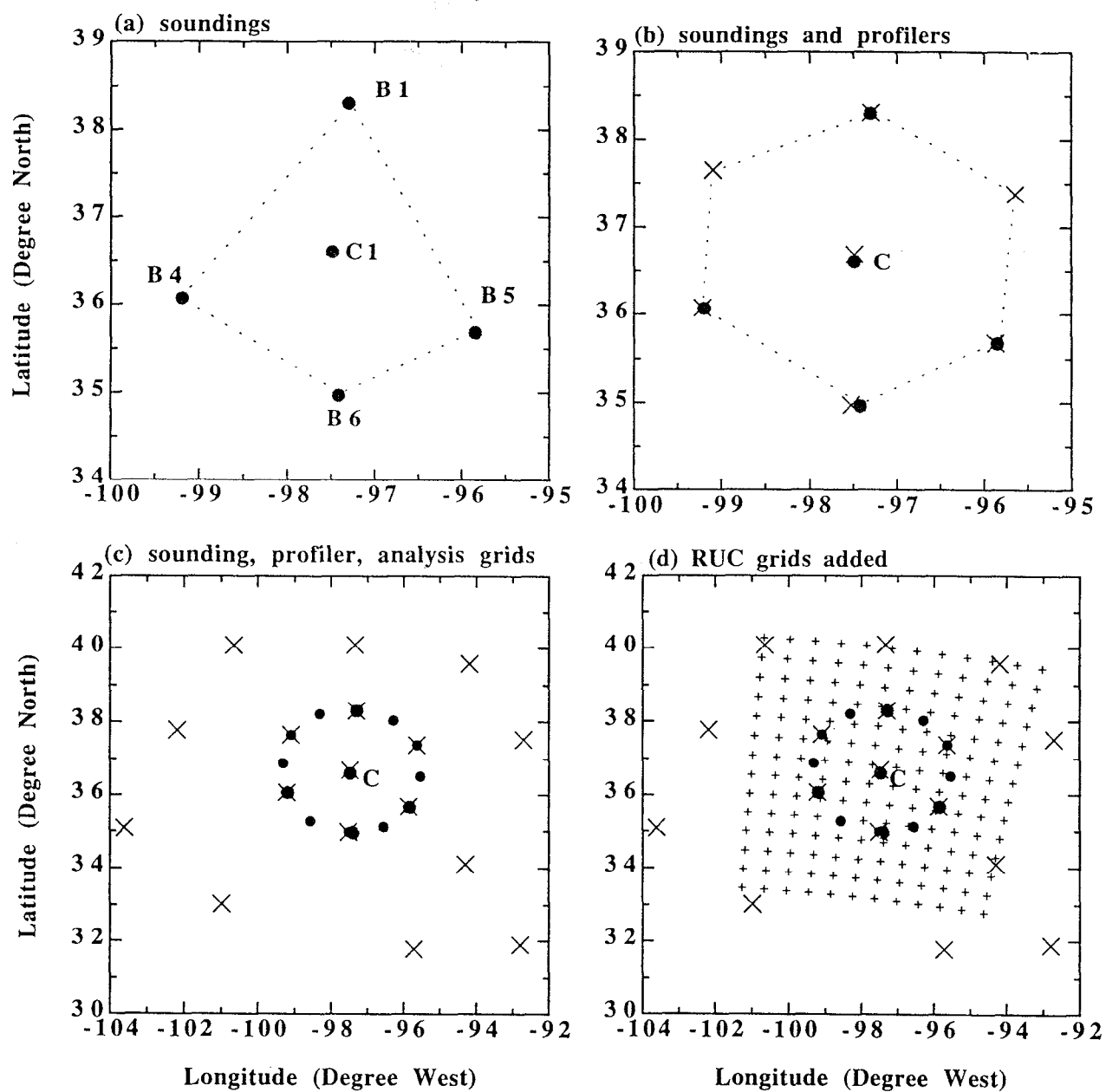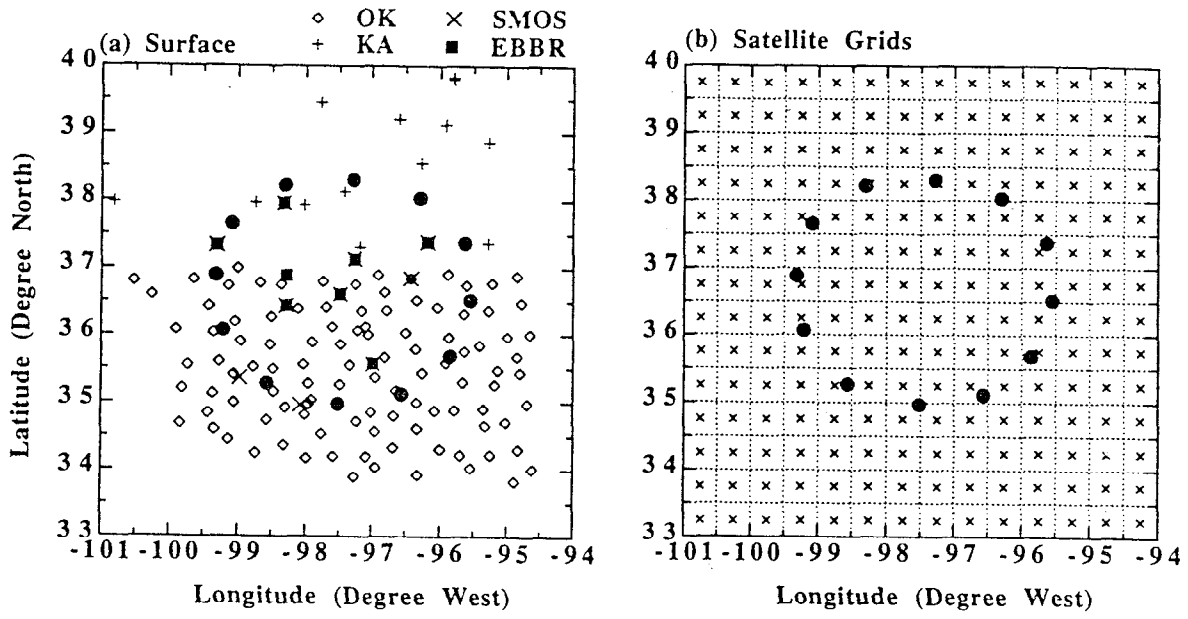
Figure 3 : Interpolated u

Figure 4 : Analysis grids and observational stations.

Figure 5 : Surface and satellite observations
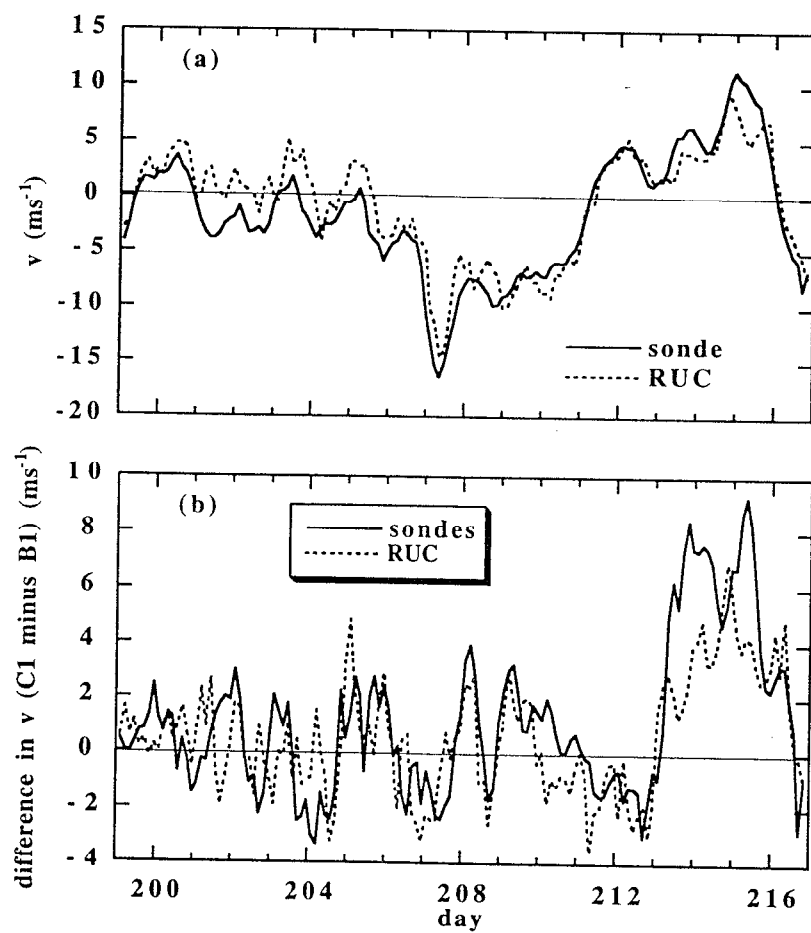
Figure 6 : v wind at 500 mb in the July 1995 ARM IOP.

500 mb v at C1
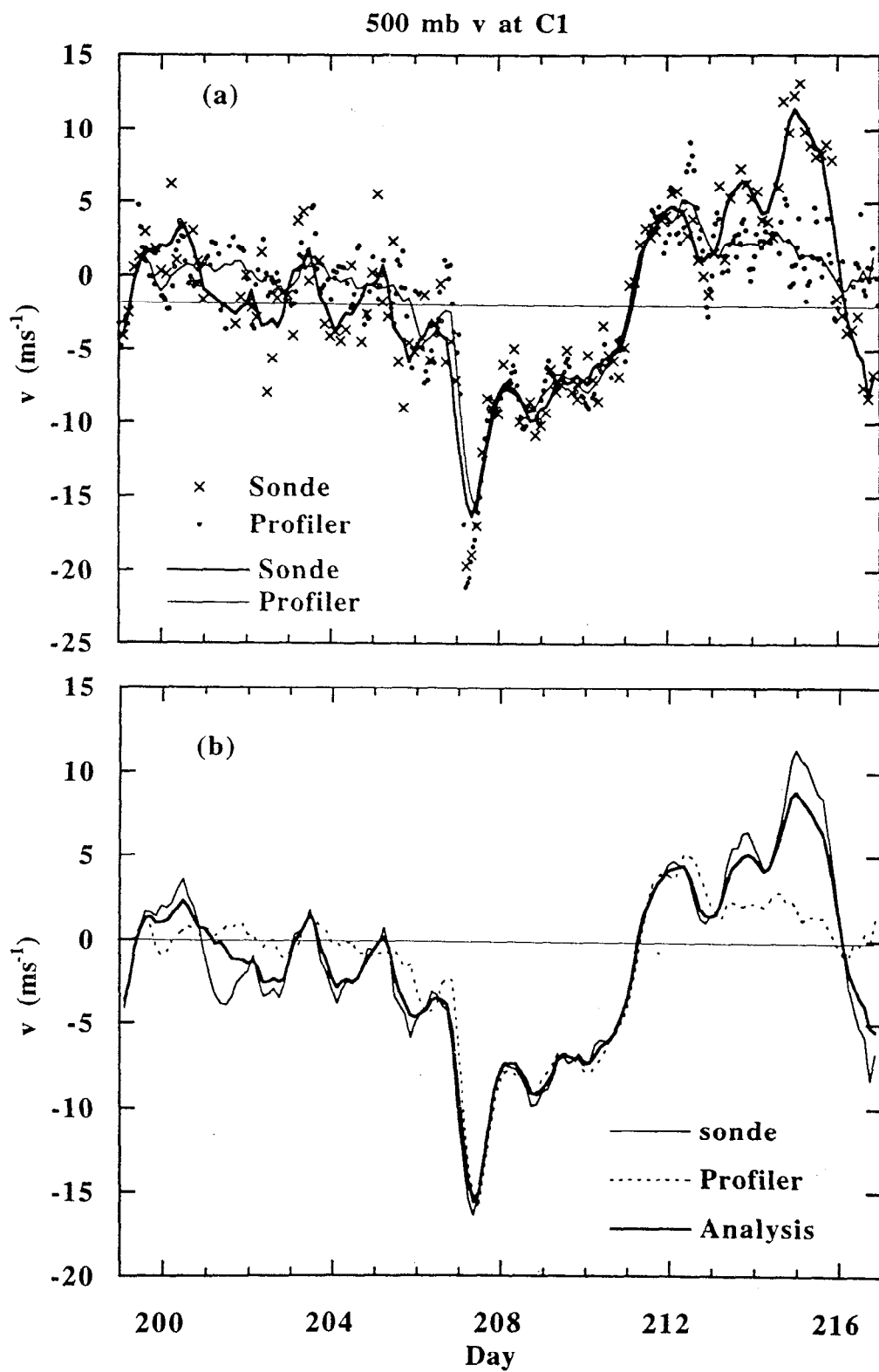
Figure 7 : v wind at 500 mb in the July 1995 ARM IOP.
(a) profiler and sondes. (b) analysis.

(a) Missing soundings among five stations

(b) Missing profiler data among seven profilers
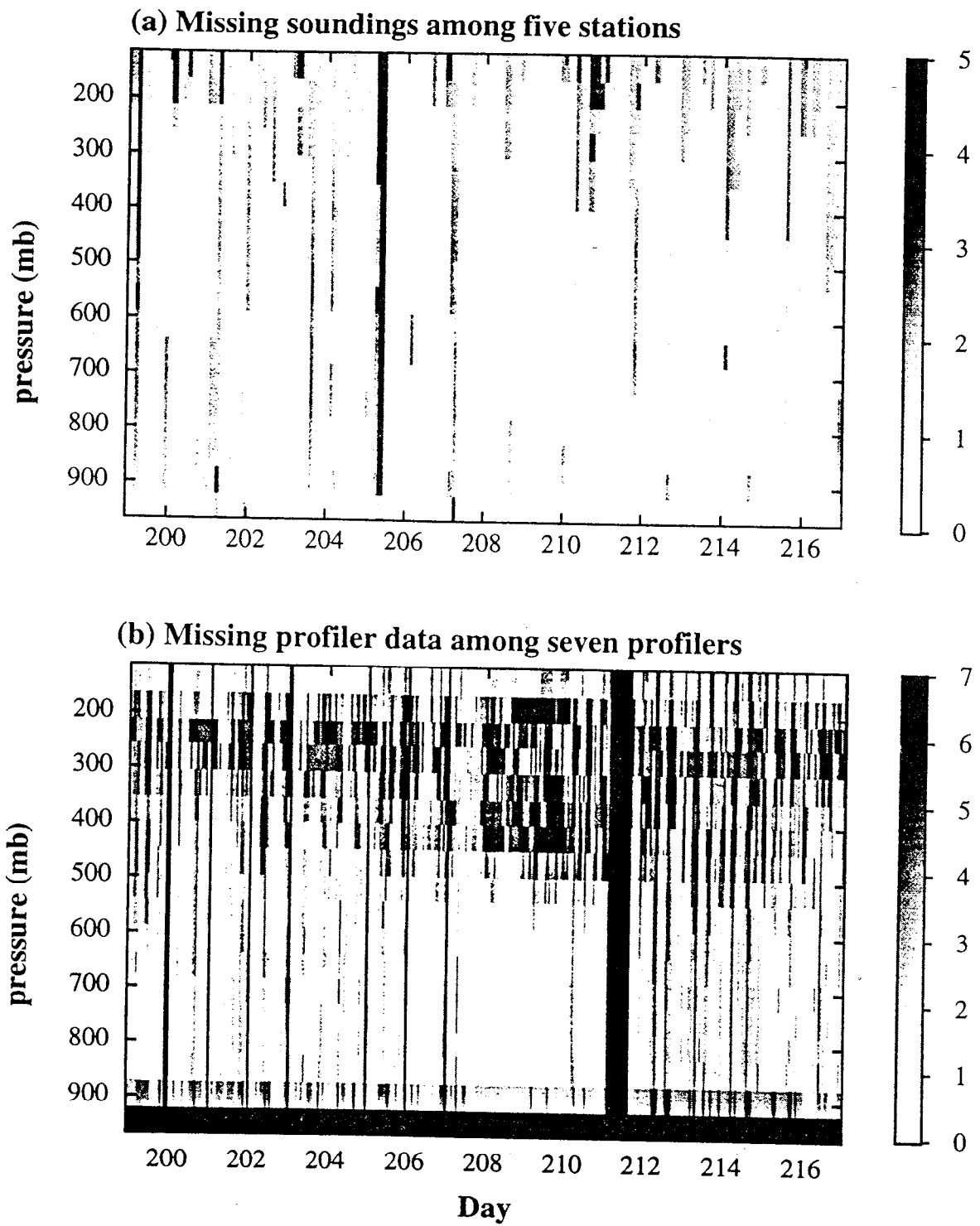
Figure 8 : Missing upper air data in the July 1995 IOP
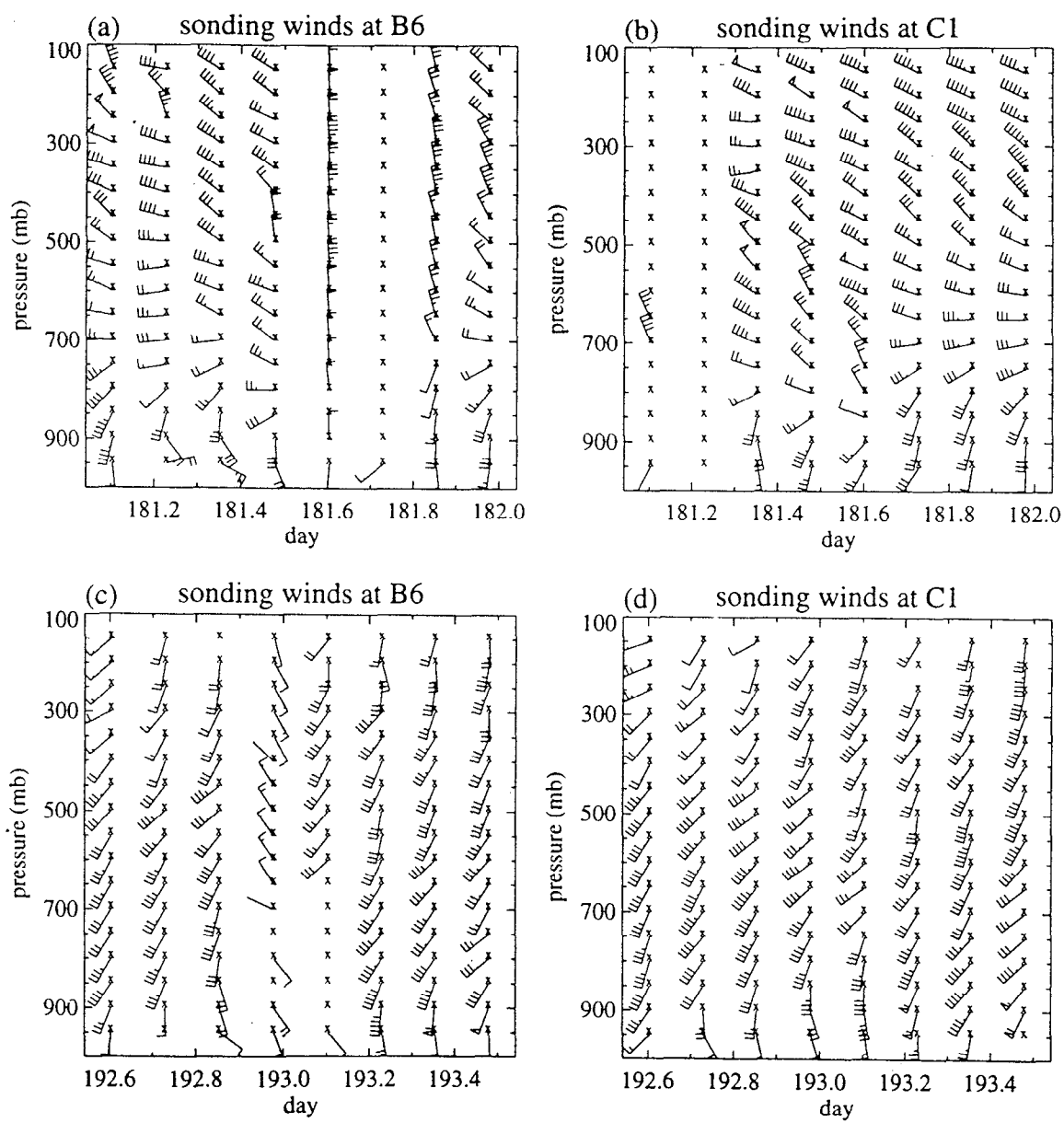
Figure 9 : winds from sondes in July 1997
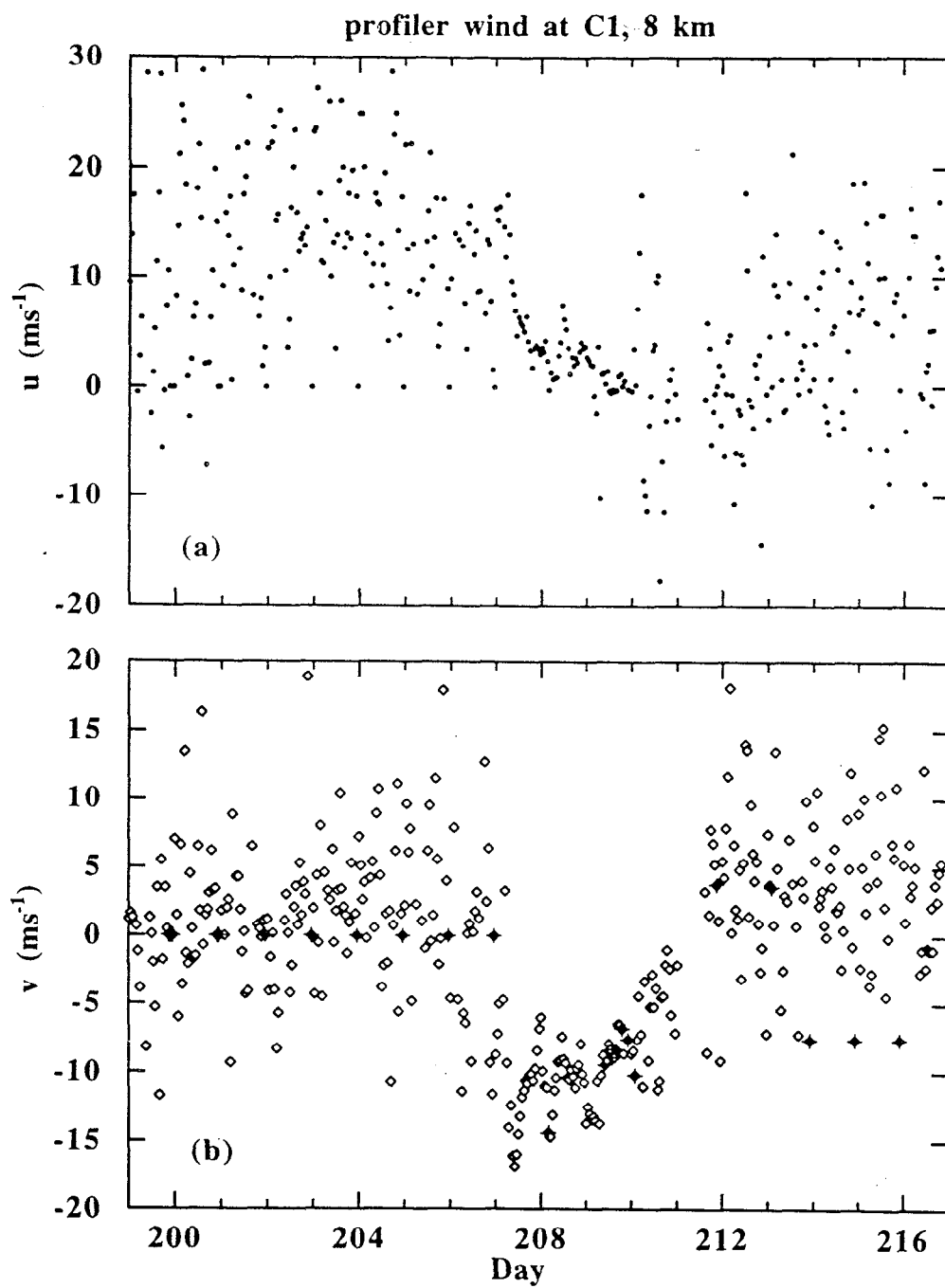
Figure 10: profiler wind at 8000 m in the July 1995 ARM IOP at C1.
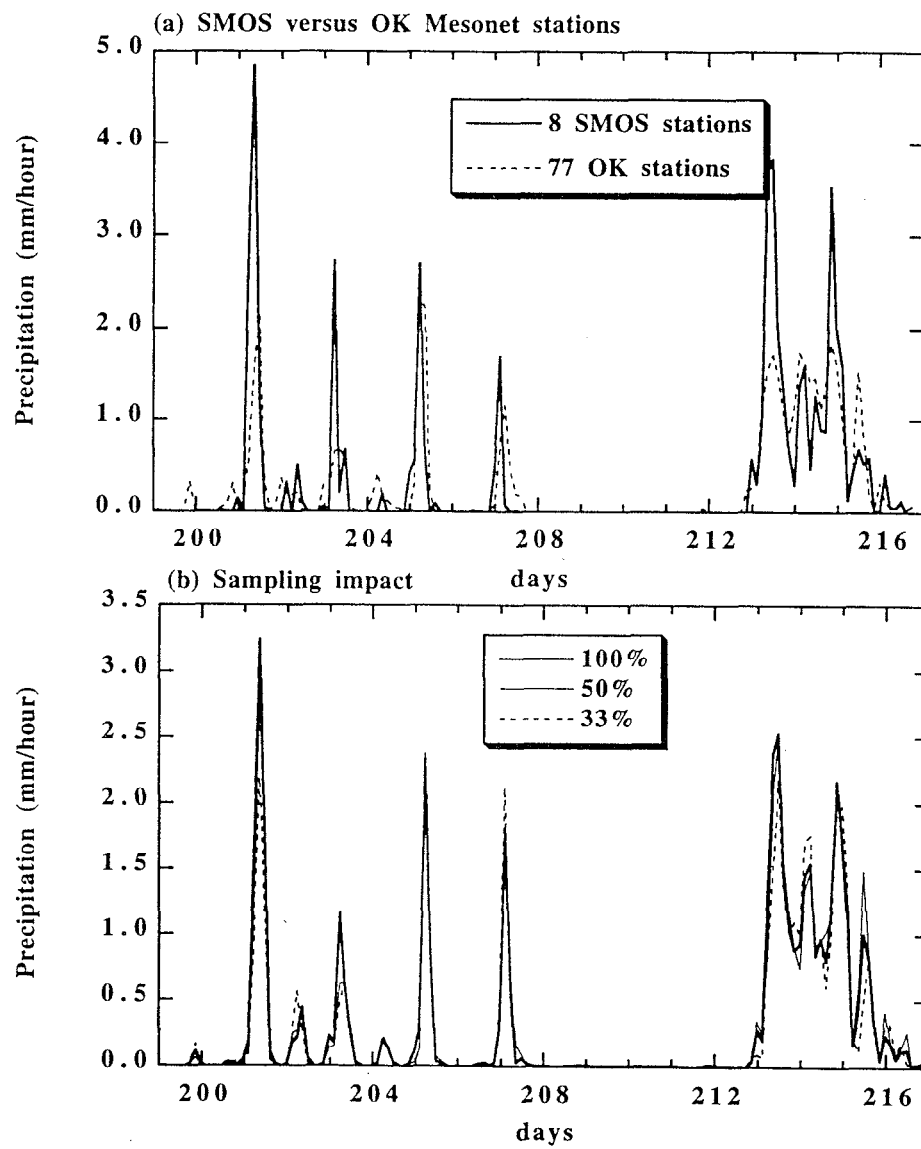
Figure II : Impact of sampled stations on mean precipitation
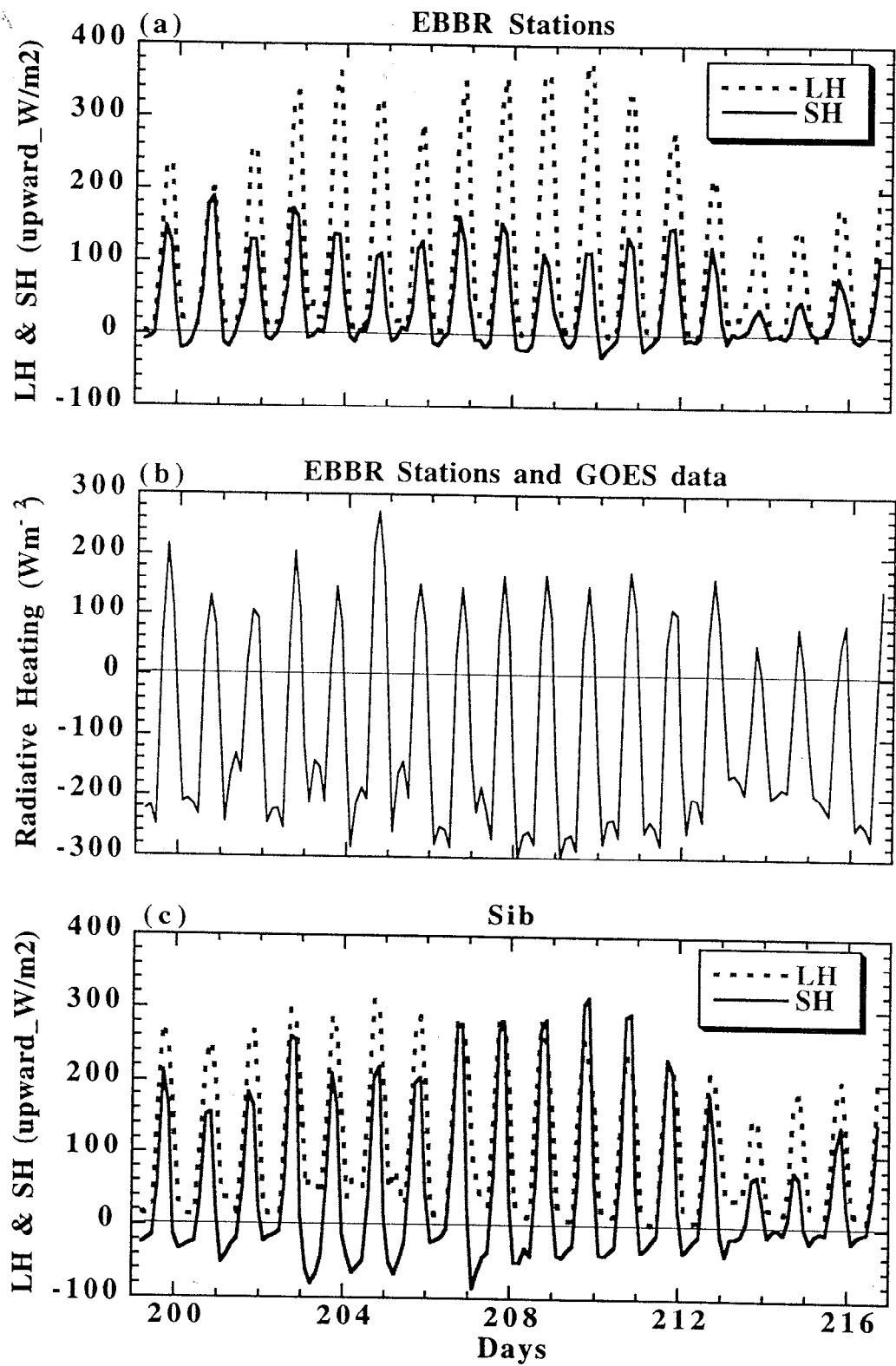
Figure 12: (a) Surface latent and sensible heat fluxes from EBBR stations.
(b) Column radiative heating. (c) Same as (a) except from Sib model.

Figure 13 : Standard deviations

(a) Temperature (K)

(b) water vapor mixing ration (g kg^{-1})

(c) wind divergence (1.0e-5 s^{-1})

(unit: 1.0e-5 s^{-1})

(d) vertical velocity (mb hour^{-1})

(unit: mb hour^{-1})
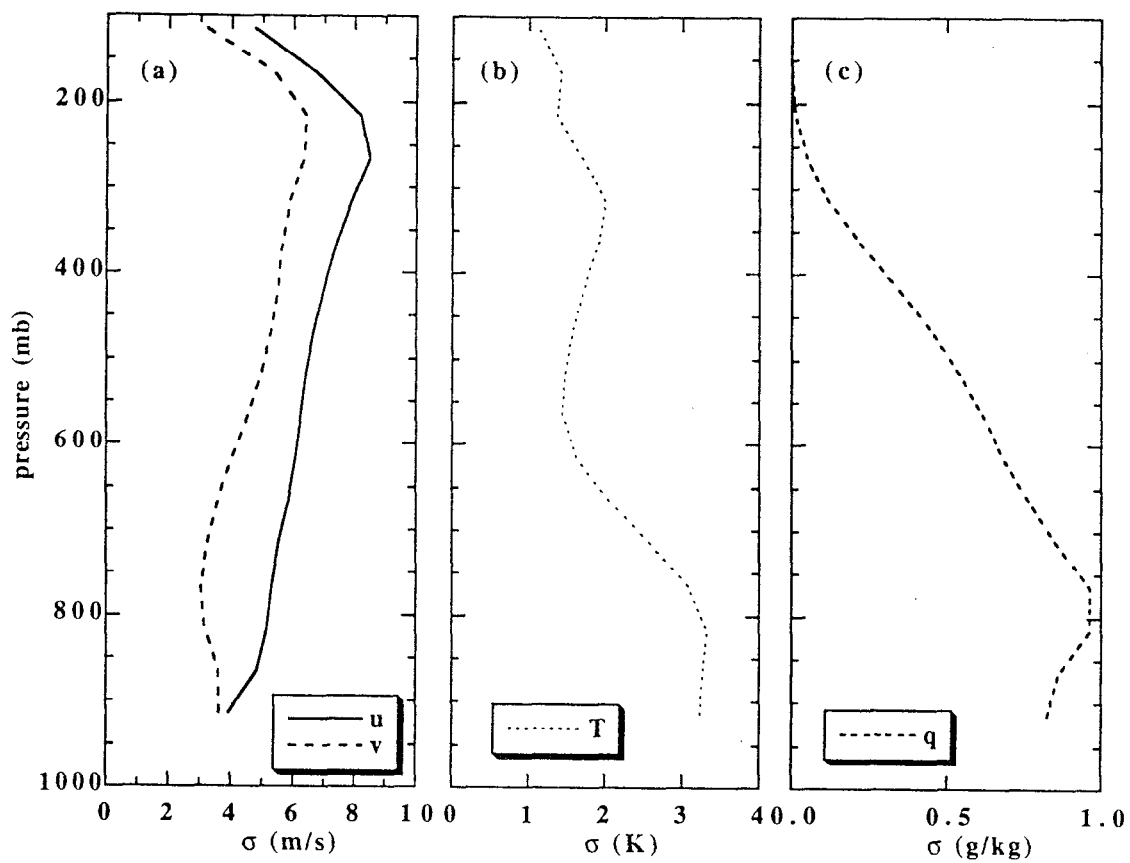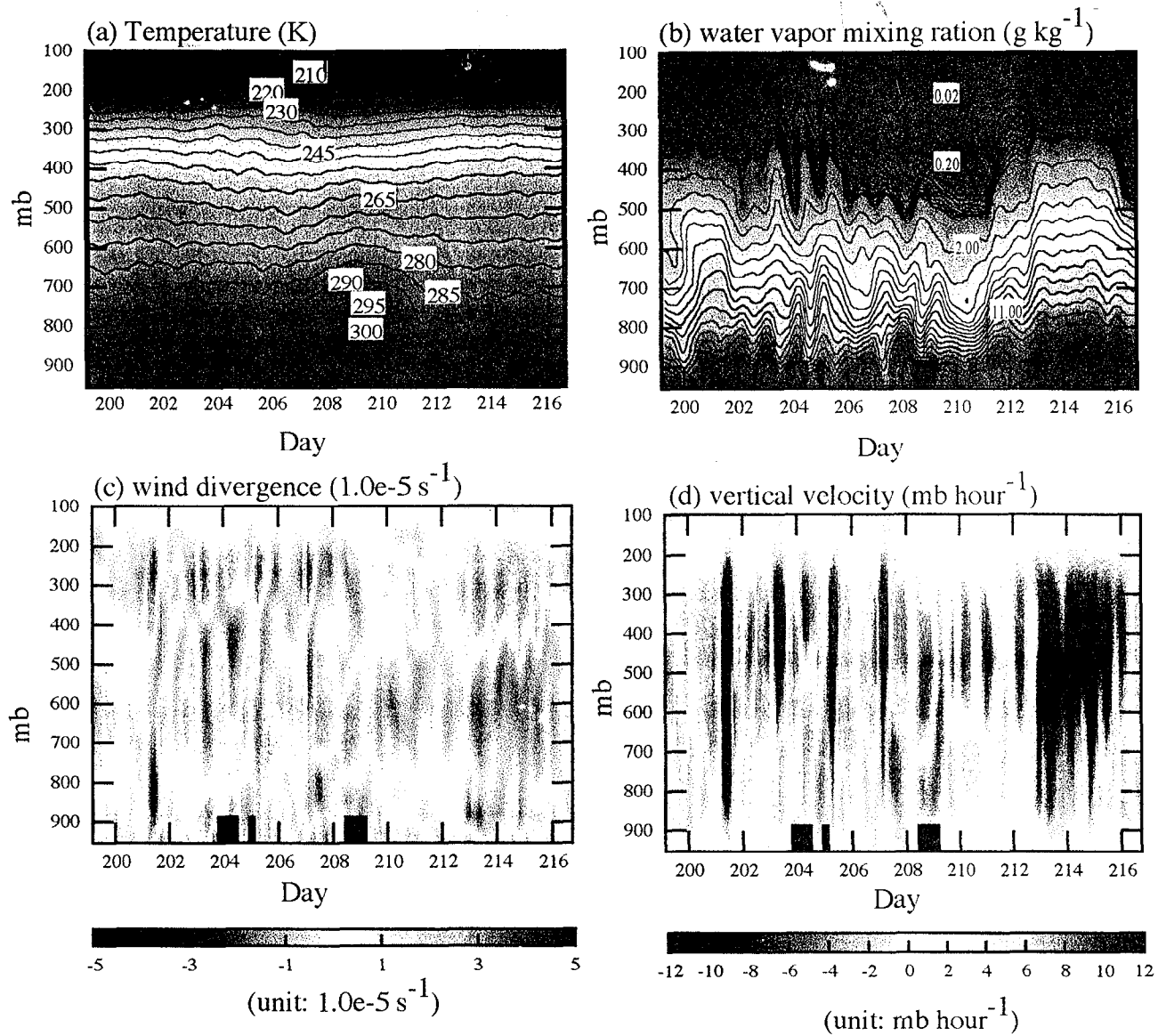
Figure 14: Analysis of the July 1995 ARM IOP data
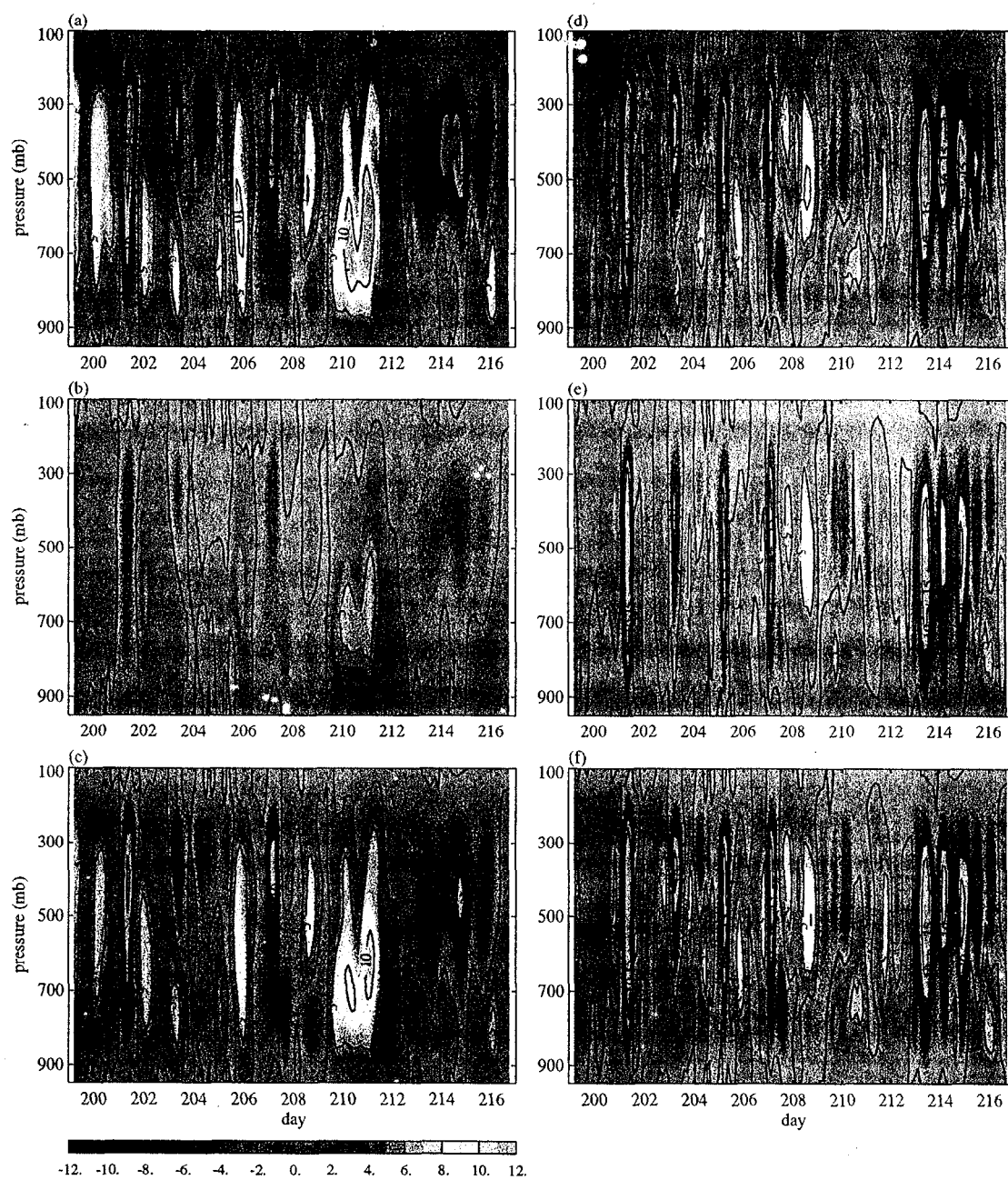
Figure 15: omega (mb/hour) for Barnes scheme
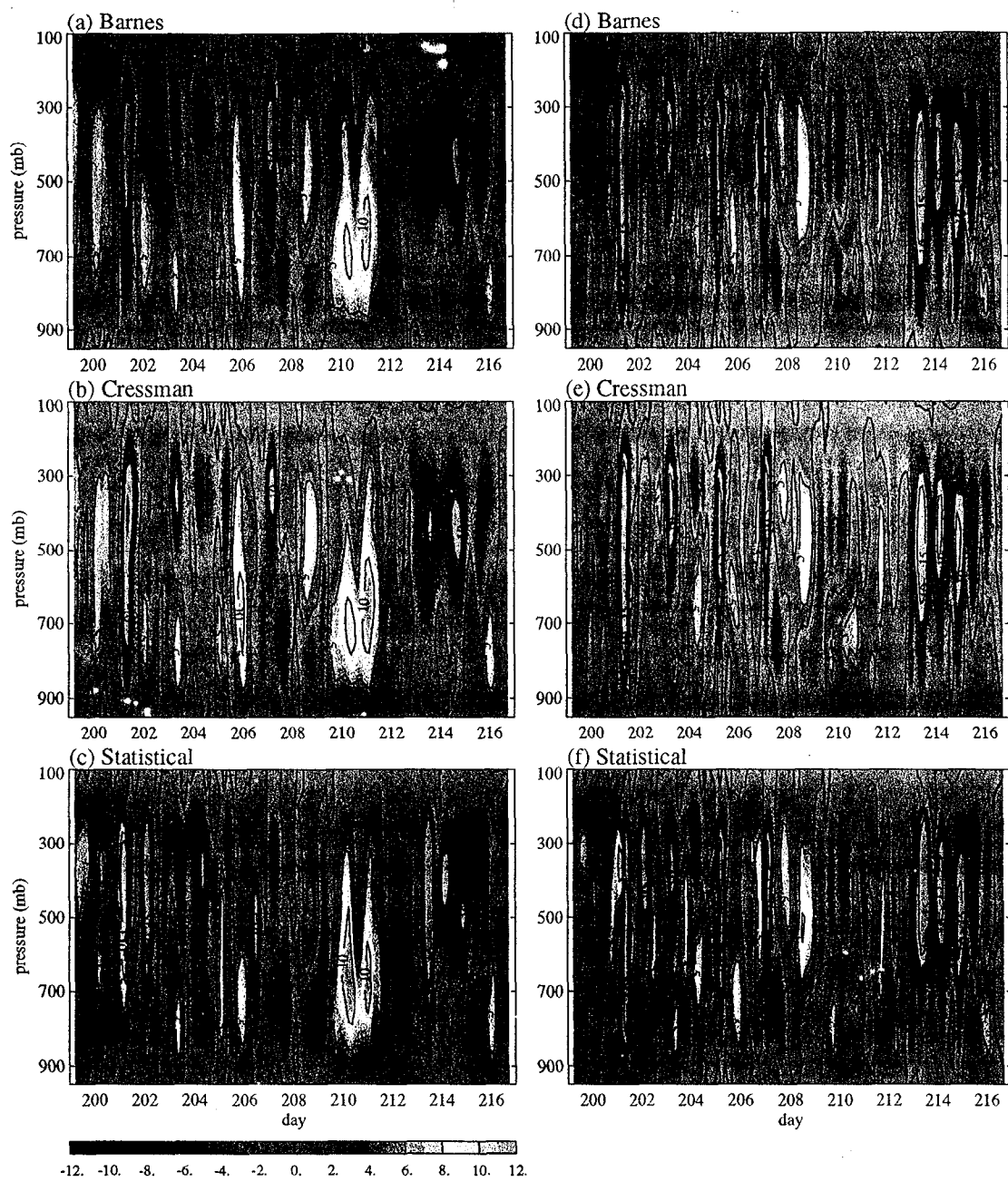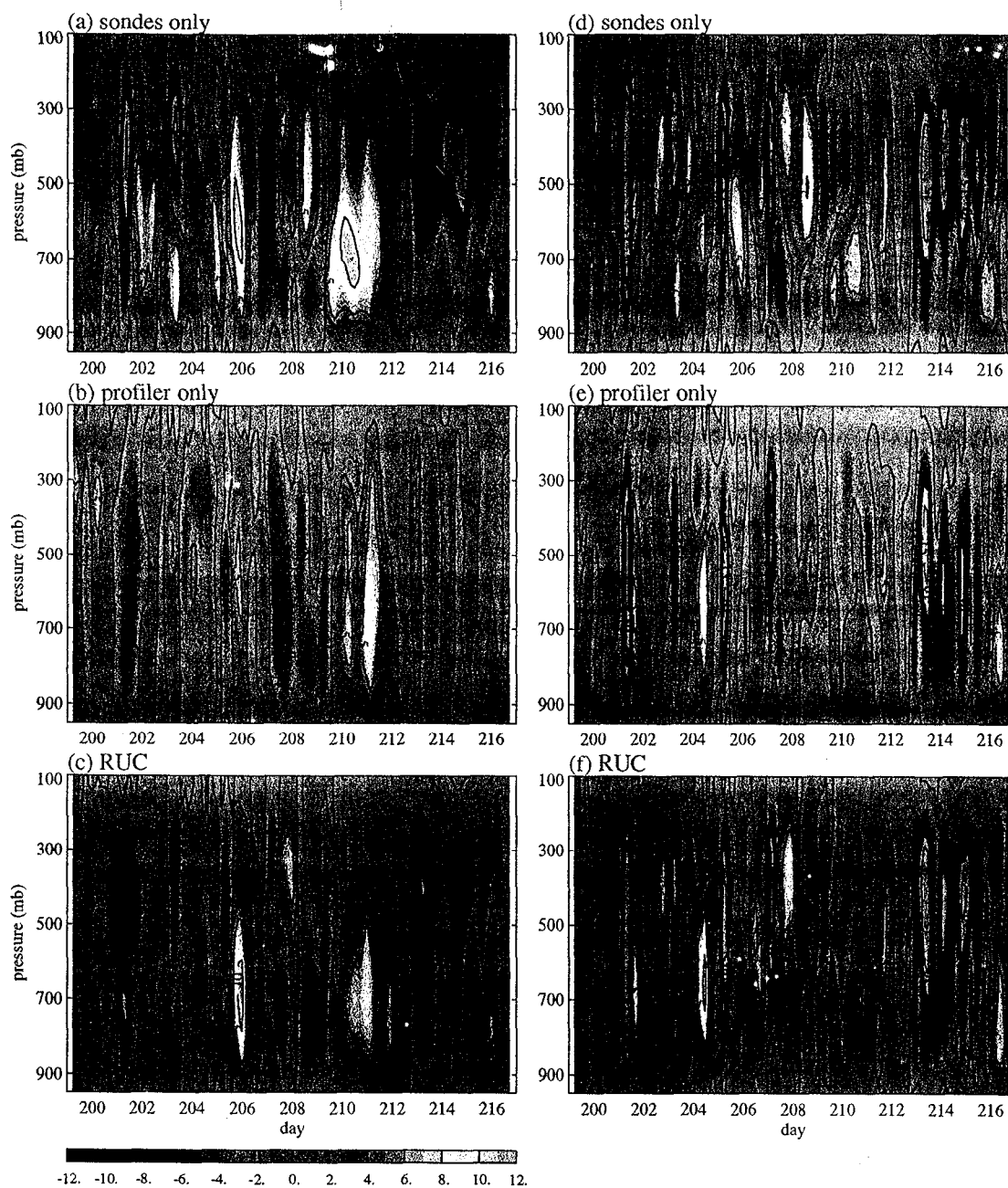
Figure 16: omega (mb/hour) from different schemes
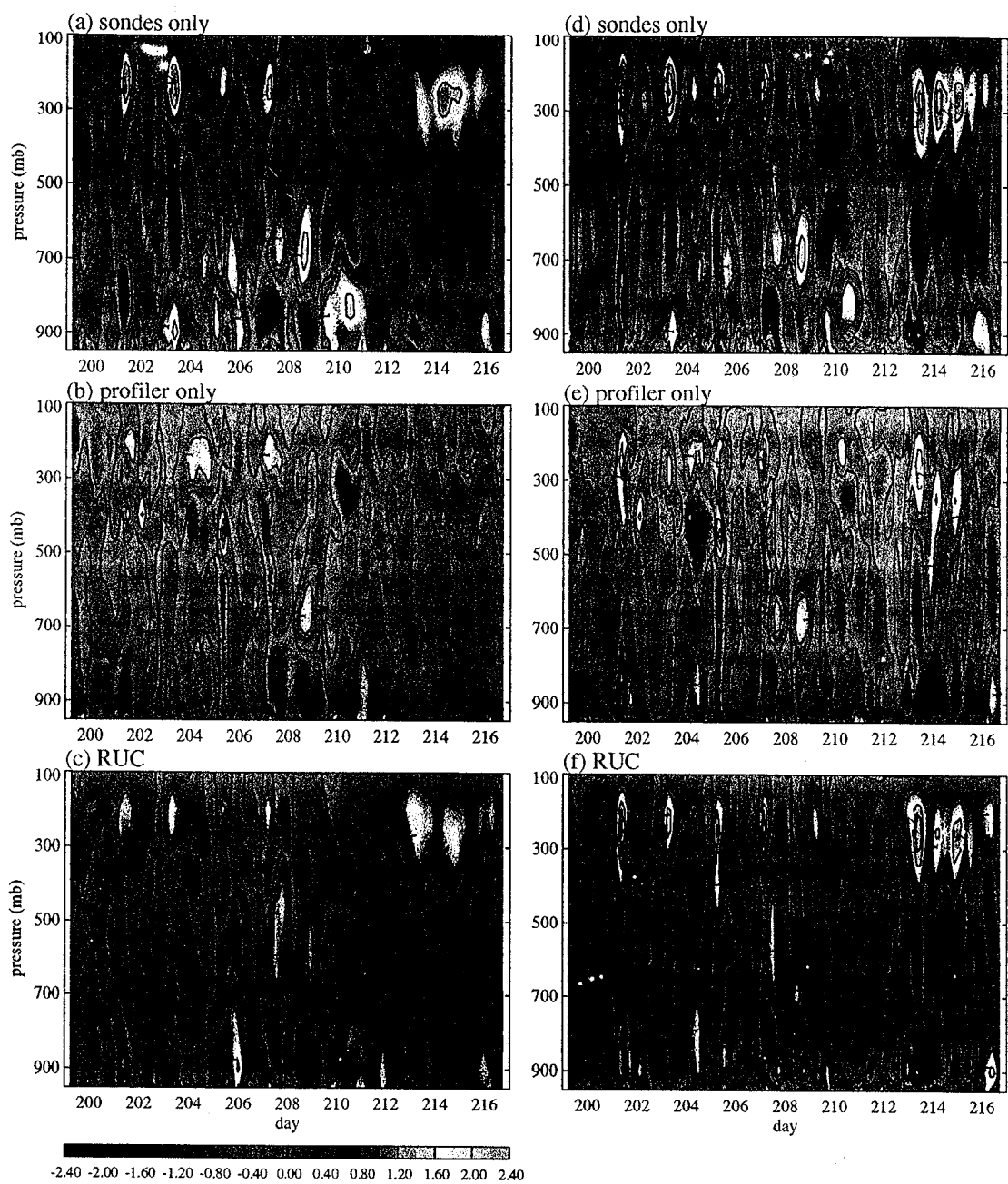
Figure 17 : omega (mb/hour) from different data sources

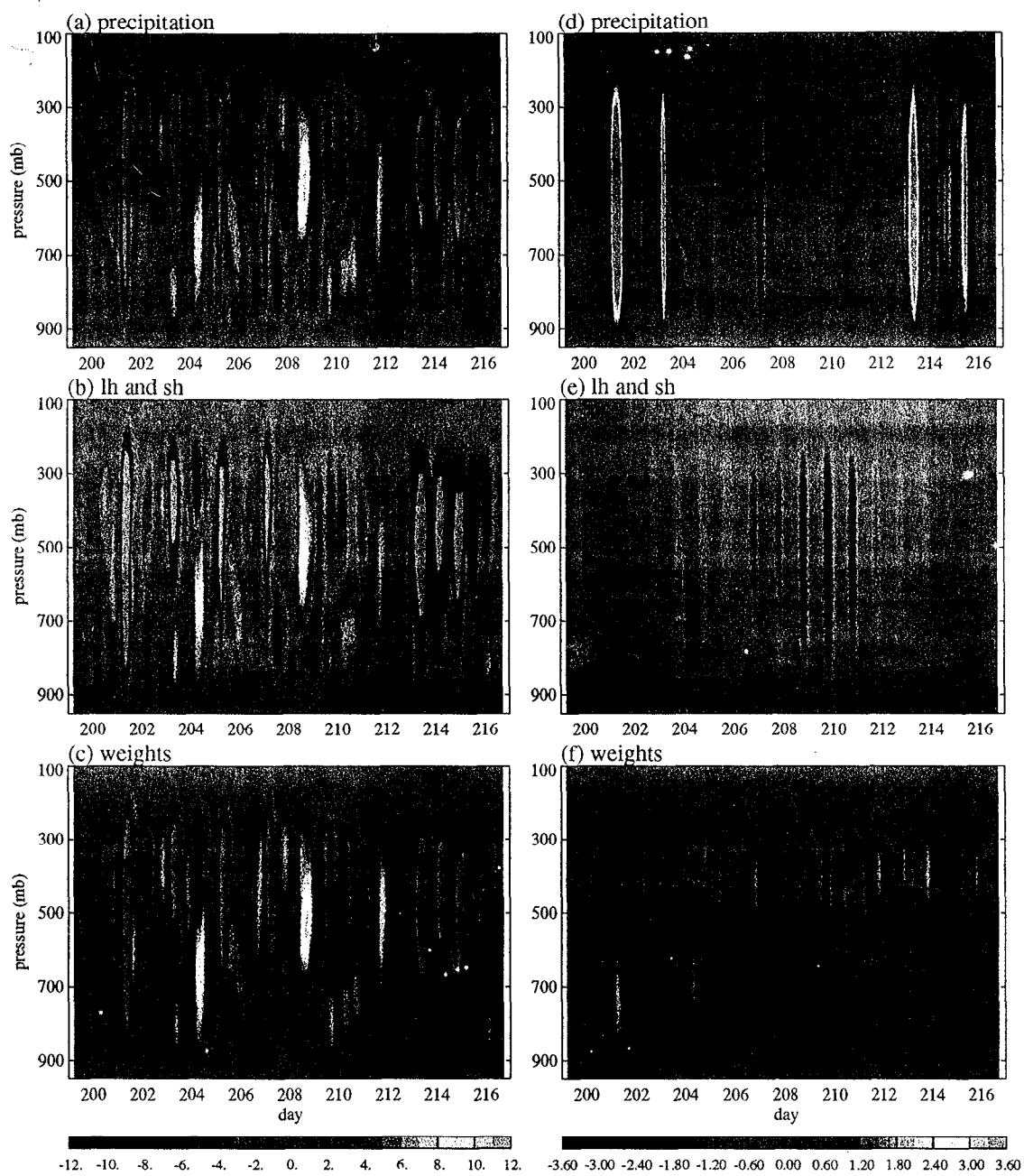Figure 18 : div (1.0E-5 1/s) from different data sources

Figure 19: omega(mb/hour) and differences for various constraint variables